GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: May 4, 2006, 06:06:17 ; Search time 39 Seconds
(without alignments)
19.737 Million cell updates/sec

Title: US-09-726-470A-2
Perfect score: 15
Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15	100.0	22	2	S47206	T-cell receptor J-
2	15	100.0	23	2	S47192	T-cell receptor J-
3	15	100.0	39	2	B85990	hypothetical prote
4	15	100.0	42	2	T07581	hypothetical prote
5	15	100.0	57	2	S16587	hypothetical prote
6	15	100.0	60	2	D90971	hypothetical prote
7	15	100.0	60	2	D85744	unknown protein en
8	15	100.0	63	2	AG1332	hypothetical prote
9	15	100.0	63	2	AG1703	hypothetical prote
10	15	100.0	77	2	JC5052	hypothetical 8.9k
11	15	100.0	79	2	B83400	hypothetical prote
12	15	100.0	79	2	I38991	tumorigenic conver
13	15	100.0	80	2	T39148	hypothetical prote
14	15	100.0	80	2	B43259	H+-transporting tw
15	15	100.0	86	2	S35769	T-cell receptor al
16	15	100.0	93	2	AC0592	probable membrane
17	15	100.0	97	2	B90725	hypothetical prote
18	15	100.0	97	2	C85576	hypothetical prote
19	15	100.0	97	2	F64809	ybgE protein - Esc
20	15	100.0	98	2	A95329	probable fragment
21	15	100.0	100	2	E69846	hypothetical prote
22	15	100.0	101	2	S51384	hypothetical prote
23	15	100.0	101	2	AE1386	transcription regu
24	15	100.0	101	2	AG1761	transcription regu
25	15	100.0	104	2	A70665	probable ureB prot
26	15	100.0	109	2	G64609	hypothetical prote
27	15	100.0	110	2	C55228	hypothetical prote
28	15	100.0	113	2	S55533	Ig heavy chain V r
29	15	100.0	113	2	S55530	Ig heavy chain V r
30	15	100.0	113	2	S55528	Ig heavy chain V r
31	15	100.0	113	2	S55532	Ig heavy chain V r
32	15	100.0	113	2	S55531	Ig heavy chain V r
33	15	100.0	113	2	F72687	hypothetical prote
34	15	100.0	114	2	D71048	hypothetical prote
35	15	100.0	116	2	G82537	pterin-4-alpha-car
36	15	100.0	119	1	CUPSAM	amicyanin precurso
37	15	100.0	119	2	E72714	probable ribosomal
38	15	100.0	121	1	G1HUHE	Ig heavy chain V-I
39	15	100.0	122	2	H82231	hypothetical prote
40	15	100.0	128	2	E70547	hypothetical prote
41	15	100.0	129	2	T12924	conserved hypothet
42	15	100.0	129	2	H72627	hypothetical prote
43	15	100.0	133	2	D48776	polyprotein (E2/NS
44	15	100.0	133	2	S15825	agglutinin alpha c
45	15	100.0	133	2	B30242	stem cell protein
46	15	100.0	135	2	AF3551	hypothetical prote
47	15	100.0	138	2	T35211	probable lyase - S
48	15	100.0	138	2	A87192	conserved hypothet
49	15	100.0	139	2	G97080	hypothetical prote
50	15	100.0	140	2	H71903	hypothetical prote
51	15	100.0	141	2	S30832	hypothetical prote
52	15	100.0	146	2	I64011	hypothetical prote
53	15	100.0	147	2	S49526	protein kinase hom
54	15	100.0	148	2	A99268	conserved hypothet
55	15	100.0	149	2	D75019	hypothetical prote
56	15	100.0	149	2	A58801	mannose-specific l
57	15	100.0	155	2	C69002	conserved hypothet
58	15	100.0	155	2	B28083	glutathione transf
59	15	100.0	155	2	B83771	hypothetical prote
60	15	100.0	157	2	H72751	hypothetical prote
61	15	100.0	157	2	E84997	16 kD heat shock p
62	15	100.0	157	2	D84063	hypothetical prote
63	15	100.0	158	2	E72702	hypothetical prote
64	15	100.0	159	2	I84615	coagulation factor
65	15	100.0	160	2	D70524	hypothetical prote
66	15	100.0	161	2	B71687	hypothetical prote
67	15	100.0	162	2	AE0446	rod shape-determin
68	15	100.0	163	2	T48289	hypothetical prote
69	15	100.0	164	2	C90442	conserved hypothet
70	15	100.0	168	2	F71503	probable glycerol-
71	15	100.0	173	2	C84368	hypothetical prote
72	15	100.0	176	2	I56314	interferon-alpha -
73	15	100.0	177	2	F83410	conserved hypothet
74	15	100.0	178	2	B84636	probable auxin-reg
75	15	100.0	180	2	T44944	hypothetical prote
76	15	100.0	181	2	S22793	partitioning prote
77	15	100.0	184	2	B86192	hypothetical prote
78	15	100.0	189	1	IVHU14	interferon alpha-I
79	15	100.0	189	1	IVHU4B	interferon alpha-I
80	15	100.0	189	1	IVHUA5	interferon alpha-5
81	15	100.0	189	1	IVHUA9	interferon alpha-1
82	15	100.0	189	2	I51970	interferon precurs
83	15	100.0	189	2	I52347	interferon alpha-M
84	15	100.0	189	2	T04499	auxin-induced prot
85	15	100.0	189	2	D71073	hypothetical prote
86	15	100.0	189	2	AC2854	acetyltransferase
87	15	100.0	192	2	S15661	(2'-5')oligo(A) sy
88	15	100.0	193	2	D85769	hypothetical prote
89	15	100.0	193	2	AC0692	probable membrane
90	15	100.0	193	2	E64919	probable membrane
91	15	100.0	193	2	AH0273	probable membrane
92	15	100.0	193	2	H90920	probable membrane
93	15	100.0	195	2	E82857	PilX protein XF003
94	15	100.0	196	2	F83312	probable type II s
95	15	100.0	197	2	G97015	imidazoleglycerol-
96	15	100.0	197	2	D82783	hypothetical prote
97	15	100.0	197	2	H85065	hypothetical prote
98	15	100.0	198	2	T04953	hypothetical prote
99	15	100.0	198	2	D97301	probable membrane
100	15	100.0	198	2	C33465	lic-1 protein C -

ALIGNMENTS

RESULT 1
S47206
T-cell receptor J-alpha wnVII.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-Jul-1999
C;Accession: S47206
R;Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A;Reference number: S40133
A;Accession: S47206
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-22 <PLA>
A;Cross-references: UNIPARC:UPI0000116127; EMBL:X71036; NID:g507043; PIDN:CAA50353.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 100.0%; Score 15; DB 2; Length 22;
Best Local Similarity 60.0%; Pred. No. 88;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 8 RALTF 12

RESULT 2
S47192
T-cell receptor J-alpha wnVII.2 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-Jul-1999
C;Accession: S47192
R;Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A;Reference number: S40133
A;Accession: S47192
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-23 <PLA>
A;Cross-references: UNIPARC:UPI0000116136; EMBL:X71051; NID:g506974; PIDN:CAA50368.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 100.0%; Score 15; DB 2; Length 23;
Best Local Similarity 60.0%; Pred. No. 92;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 9 RALTF 13

RESULT 3
B85990
hypothetical protein Z4614 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85990
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85990
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <STO>
A;Cross-references: UNIPROT:Q8X417; UNIPARC:UPI00000D0E65; GB:AE005174; NID:g12517881; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z4614

Query Match 100.0%; Score 15; DB 2; Length 39;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 28 RALAF 32

RESULT 4
T07581
hypothetical protein 42g - Japanese black pine chloroplast
C;Species: chloroplast Pinus thunbergiana (Japanese black pine)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C;Accession: T07581
R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A;Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genom
A;Reference number: Z16030; MUID:95024047; PMID:7937893
A;Accession: T07581
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-42 <WAK>
A;Cross-references: UNIPROT:Q33005; UNIPARC:UPI00000939E4; EMBL:D17510; NID:g529643; PI
C;Genetics:
A;Genome: chloroplast
C;Keywords: chloroplast

Query Match 100.0%; Score 15; DB 2; Length 42;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RSLSF 18

RESULT 5
S16587
hypothetical protein 1 - lamb's-quarters
C;Species: Chenopodium album (lamb's-quarters)
C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 26-May-1995
C;Accession: S16587
R;Doerfel, P.; Weihe, A.; Dolferus, R.; Boerner, T.
Plant Mol. Biol. 17, 155-156, 1991
A;Title: DNA sequence of a mitochondrial plasmid from Chenopodium album.
A;Reference number: S16587; MUID:91329724; PMID:1651127
A;Accession: S16587
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-57 <DOE>
A;Cross-references: UNIPARC:UPI000017AF1D; EMBL:X58911

Query Match 100.0%; Score 15; DB 2; Length 57;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 26 RTLTF 30

RESULT 6
D90971
hypothetical protein ECs2740 [imported] - Escherichia coli (strain O157:H7, substrain R
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: D90971
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90971
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-60 <HAY>
A;Cross-references: UNIPROT:Q8X8T7; UNIPARC:UPI00000D08D0; GB:BA000007; PIDN:BAB36163.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2740

Query Match 100.0%; Score 15; DB 2; Length 60;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 29 RALAF 33

RESULT 7
D85744
unknown protein encoded within prophage CP-933R [imported] - Escherichia coli (strain O1
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85744
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85744
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-60 <STO>
A;Cross-references: UNIPROT:Q8X8T7; UNIPARC:UPI00000D08D0; GB:AE005174; NID:g12515365; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z2370

Query Match 100.0%; Score 15; DB 2; Length 60;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 29 RALAF 33

RESULT 8
AG1332
hypothetical protein lmo2063 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1332
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1332
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <GLA>
A;Cross-references: UNIPROT:Q8Y5J7; UNIPARC:UPI00000555B3; GB:NC_003210; PIDN:CAD00141.1
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2063

Query Match 100.0%; Score 15; DB 2; Length 63;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 39 RTLTF 43

RESULT 9
AG1703
hypothetical protein lin2169 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1703
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeck
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1703
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <GLA>
A;Cross-references: UNIPROT:Q929V4; UNIPARC:UPI00000CC793; GB:AL592022; PIDN:CAC97399.
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2169

Query Match 100.0%; Score 15; DB 2; Length 63;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 39 RTLTF 43

RESULT 10
JC5052
hypothetical 8.9k protein - Shigella flexneri
C;Species: Shigella flexneri
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C;Accession: JC5052
R;Venkatesan, M.M.; Alexander, W.A.; Fernandez-Prada, C.
Gene 175, 23-27, 1996
A;Title: A Shigella flexneri invasion plasmid gene, ipgH, with homology to IS629 and s
A;Reference number: JC5050; MUID:97074644; PMID:8917071
A;Accession: JC5052
A;Molecule type: DNA
A;Residues: 1-77 <VEN>
A;Cross-references: UNIPROT:Q54148; UNIPARC:UPI00000B43C1; GB:U28354; NID:g1016674; PI
A;Note: in the authors' translation, residues 5-7 are shown after residue 15, residues
C;Superfamily: DNA replication protein dnaC

Query Match 100.0%; Score 15; DB 2; Length 77;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 49 RSLSF 53

RESULT 11
B83400
hypothetical protein PA1970 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: B83400
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: B83400
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-79 <STO>
A;Cross-references: UNIPROT:Q9I2D4; UNIPARC:UPI00000C5538; GB:AE004623; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1970

Query Match 100.0%; Score 15; DB 2; Length 79;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 44 RSLSF 48

RESULT 12
I38991
tumorigenic conversion-associated protein CATR1 - human
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C;Accession: I38991
R;Li, D.; Noyes, I.; Shuler, C.; Milo, G.E.
Proc. Natl. Acad. Sci. U.S.A. 92, 6409-6413, 1995
A;Title: Cloning and sequencing of CATR1.3, a human gene associated with tumorigenic con
A;Reference number: I38991; MUID:95327656; PMID:7604004
A;Accession: I38991
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-79 <RES>
A;Cross-references: UNIPROT:Q13166; UNIPARC:UPI00001270C8; EMBL:U25433; NID:g896044; PID
C;Genetics:
A;Gene: GDB:CATR1
A;Cross-references: GDB:633071; OMIM:600676
A;Map position: 16p13.3-16p13.3

Query Match 100.0%; Score 15; DB 2; Length 79;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 42 RALTF 46

RESULT 13
T39148
hypothetical protein SPAC8C9.11 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39148
R;Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21748
A;Accession: T39148
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-80 <OLI>
A;Cross-references: UNIPROT:O14280; UNIPARC:UPI000006AC4F; EMBL:Z99168; PIDN:CAB16299.1;
A;Experimental source: strain 972h-; cosmid c8C9
C;Genetics:
A;Gene: SPDB:SPAC8C9.11
A;Map position: 1
A;Introns: 20/3

Query Match 100.0%; Score 15; DB 2; Length 80;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 72 RTLSF 76

RESULT 14
B43259
H+-transporting two-sector ATPase (EC 3.6.3.14) chain a - Enterococcus hirae (fragment)
C;Species: Enterococcus hirae
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B43259
R;Shibata, C.; Ehara, T.; Tomura, K.; Igarashi, K.; Kobayashi, H.
J. Bacteriol. 174, 6117-6124, 1992
A;Title: Gene structure of Enterococcus hirae (Streptococcus faecalis) F1F0-ATPase, whi
A;Reference number: A43259; MUID:93015650; PMID:1328152
A;Accession: B43259
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-80 <SHI>
A;Cross-references: UNIPROT:P43454; UNIPARC:UPI000017CD83
A;Experimental source: ATCC 9790
A;Note: sequence extracted from NCBI backbone (NCBIN:115116, NCBIP:115124)
C;Keywords: hydrolase

Query Match 100.0%; Score 15; DB 2; Length 80;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 4 RSLTF 8

RESULT 15
S35769
T-cell receptor alpha chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C;Accession: S35769
R;Wedderburn, L.R.
submitted to the EMBL Data Library, June 1993
A;Reference number: S35769
A;Accession: S35769
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-86 <WED>
A;Cross-references: UNIPARC:UPI0000116542; EMBL:Z22965; NID:g312153; PIDN:CAA80538.1; P
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 100.0%; Score 15; DB 2; Length 86;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 56 RALTF 60

RESULT 16
AC0592
probable membrane protein STY0789 [imported] - Salmonella enterica subsp. enterica sero
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0592
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0592
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-93 <PAR>
A;Cross-references: UNIPARC:UPI000005A119; GB:AL513382; PIDN:CAD05205.1; PID:g16501975;
C;Genetics:
A;Gene: STY0789

Query Match 100.0%; Score 15; DB 2; Length 93;
Best Local Similarity 60.0%; Pred. No. 3.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RALSF 18

RESULT 17
B90725
hypothetical protein ECs0770 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B90725
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-97 <HAY>
A;Cross-references: UNIPROT:P37343; UNIPARC:UPI000013A331; GB:BA000007; PIDN:BAB34193.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0770

Query Match 100.0%; Score 15; DB 2; Length 97;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 18 RALSF 22

RESULT 18
C85576
hypothetical protein ybgE [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: C85576
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-97 <STO>
A;Cross-references: UNIPROT:P37343; UNIPARC:UPI000013A331; GB:AE005174; NID:g12513668; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ybgE

Query Match 100.0%; Score 15; DB 2; Length 97;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 18 RALSF 22

RESULT 19
F64809
ybgE protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: F64809
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F64809
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-97 <BLAT>
A;Cross-references: UNIPROT:P37343; UNIPARC:UPI000013A331; GB:AE000177; GB:U00096; NIE
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ybgE

Query Match 100.0%; Score 15; DB 2; Length 97;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 18 RALSF 22

RESULT 20
A95329
probable fragment of transposase protein [imported] - Sinorhizobium meliloti (strain 1
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: A95329
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bo
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melil
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: A95329
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <KUR>
A;Cross-references: UNIPROT:Q92ZE9; UNIPARC:UPI00000CB110; GB:AE006469; PIDN:AAK65195
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Huble
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaur
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa0997
A;Genome: plasmid

Query Match 100.0%; Score 15; DB 2; Length 98;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 81 RALTF 85

RESULT 21
E69846
hypothetical protein yjcE - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: E69846
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69846
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-100 <KUN>
A;Cross-references: UNIPROT:O31627; UNIPARC:UPI0000060231; GB:Z99110; GB:AL009126; NID:g
A;Experimental source: strain 168
C;Genetics:
A;Gene: yjcE
C;Superfamily: Bacillus subtilis hypothetical protein yjcE

  Query Match 100.0%; Score 15; DB 2; Length 100;
  Best Local Similarity 60.0%; Pred. No. 3.8e+02;
  Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 24 RALTF 28

RESULT 22
S51384
hypothetical protein YLR366w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L8039.7
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 19-Apr-2002
C;Accession: S51384
R;Du, Z.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmid 8039.
A;Reference number: S51377
A;Accession: S51384
A;Molecule type: DNA
A;Residues: 1-101 <DUZ>
A;Cross-references: UNIPARC:UPI000011E4D6; EMBL:U19103; NID:g609404; PIDN:AAB67568.1; PI
C;Genetics:
A;Gene: MIPS:YLR366w
A;Cross-references: SGD:S0004358
A;Map position: 12R
C;Superfamily: Saccharomyces hypothetical protein YLR366w

  Query Match 100.0%; Score 15; DB 2; Length 101;
  Best Local Similarity 60.0%; Pred. No. 3.8e+02;
  Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 48 RTLTF 52

RESULT 23
AE1386
transcription regulator ArsR family homolog lmo2493 [imported] - Listeria monocytogenes
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE1386
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-101 <GLA>
A;Cross-references: UNIPROT:Q8Y4F1; UNIPARC:UPI000005570B; GB:NC_003210; PIDN:CAD00571.
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2493

  Query Match 100.0%; Score 15; DB 2; Length 101;
  Best Local Similarity 60.0%; Pred. No. 3.8e+02;
  Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 47 RTLAF 51

RESULT 24
AG1761
transcription regulator ArsR family homolog lin2636 [imported] - Listeria innocua (stra
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1761
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1761
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-101 <GLA>
A;Cross-references: UNIPROT:Q928A0; UNIPARC:UPI00000CC92B; GB:AL592022; PIDN:CAC97863.1
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2636

  Query Match 100.0%; Score 15; DB 2; Length 101;
  Best Local Similarity 60.0%; Pred. No. 3.8e+02;
  Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 47 RTLTF 51

RESULT 25
A70665
probable ureB protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: A70665
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

```
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: A70665
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-104 <COL>
A;Cross-references: UNIPROT:P50048; UNIPARC:UPI0000137D88; GB:Z83859; GB:AL123456; NID:g
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: ureB
C;Superfamily: urease, beta subunit; urease 12K chain homology
F;3-100/Domain: urease 12K chain homology <U12>

  Query Match             100.0%;  Score 15;  DB 2;  Length 104;
  Best Local Similarity   60.0%;  Pred. No. 3.9e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         48 RALSF 52


RESULT 26
G64609
hypothetical protein HP0719 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: G64609
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: G64609
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-109 <TOM>
A;Cross-references: UNIPROT:O25421; UNIPARC:UPI00000C088A; GB:AE000585; GB:AE000511; NID

  Query Match             100.0%;  Score 15;  DB 2;  Length 109;
  Best Local Similarity   60.0%;  Pred. No. 4.1e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         24 RSLAF 28


RESULT 27
C55228
hypothetical protein - Thermus aquaticus (fragment)
C;Species: Thermus aquaticus
C;Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 12-Jul-2004
C;Accession: C55228
R;Meinnel, T.; Blanquet, S.
J. Bacteriol. 176, 7387-7390, 1994
A;Title: Characterization of the Thermus thermophilus locus encoding peptide deformylase
A;Reference number: A55228; MUID:95050326; PMID:7961514
A;Accession: C55228
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <MEI>
A;Cross-references: UNIPARC:UPI000013AEED; GB:X79087; NID:g602912; PIDN:CAA55697.1; PID:
C;Superfamily: P-loop kinase

  Query Match             100.0%;  Score 15;  DB 2;  Length 110;
  Best Local Similarity   60.0%;  Pred. No. 4.1e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         59 RALAF 63


RESULT 28
S55533
Ig heavy chain V region pe25 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C;Accession: S55533
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies us
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55533
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-113 <BOE>
A;Cross-references: UNIPARC:UPI0000116205; EMBL:X82594; NID:g854314; PIDN:CAA57930.1;
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             100.0%;  Score 15;  DB 2;  Length 113;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         83 RSLTF 87


RESULT 29
S55530
Ig heavy chain V region pe17 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C;Accession: S55530
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies us
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55530
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-113 <BOE>
A;Cross-references: UNIPARC:UPI00001161FD; EMBL:X82586; NID:g854296; PIDN:CAA57922.1;
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             100.0%;  Score 15;  DB 2;  Length 113;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         83 RSLTF 87


RESULT 30
S55528
Ig heavy chain V region (pe16/pe14) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C;Accession: S55528; S55529
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies us
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55528
```

```
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-113 <BOE>
A;Cross-references: UNIPARC:UPI00001161FB; EMBL:X82585; NID:g854294; PIDN:CAA57921.1; PI
A;Accession: S55529
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-113 <BO2>
A;Cross-references: UNIPARC:UPI00001161FB; EMBL:X82583; NID:g854290; PIDN:CAA57919.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             100.0%;  Score 15;  DB 2;  Length 113;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         83 RSLTF 87


RESULT 31
S55532
Ig heavy chain V region pe19 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C;Accession: S55532
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usin
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55532
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-113 <BOE>
A;Cross-references: UNIPARC:UPI00001161FF; EMBL:X82588; NID:g854300; PIDN:CAA57924.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             100.0%;  Score 15;  DB 2;  Length 113;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         83 RSLTF 87


RESULT 32
S55531
Ig heavy chain V region pe18 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000
C;Accession: S55531
R;Boettger, V.; Boettger, A.; Lane, E.B.; Spruce, B.A.
J. Mol. Biol. 247, 932-946, 1995
A;Title: Comprehensive epitope analysis of monoclonal anti-proenkephalin antibodies usin
utations in the variable region genes.
A;Reference number: S55528; MUID:95239763; PMID:7536850
A;Accession: S55531
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-113 <BOE>
A;Cross-references: UNIPARC:UPI00001161FE; EMBL:X82587; NID:g854298; PIDN:CAA57923.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;14-97/Domain: immunoglobulin homology <IMM>

  Query Match             100.0%;  Score 15;  DB 2;  Length 113;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         83 RSLTF 87


RESULT 33
F72687
hypothetical protein APE0918 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F72687
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72687
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <KAW>
A;Cross-references: UNIPROT:Q9YDJ5; UNIPARC:UPI000005DD3B; DDBJ:AP000060; NID:g5104188;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0918
C;Superfamily: Aeropyrum pernix hypothetical protein APE0918

  Query Match             100.0%;  Score 15;  DB 2;  Length 113;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         60 RSLSF 64


RESULT 34
D71048
hypothetical protein PH1676 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C;Accession: D71048
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: D71048
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-114 <KAW>
A;Cross-references: UNIPROT:O59326; UNIPARC:UPI00000631C1; GB:AP000006; NID:g3236133; P
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1676

  Query Match             100.0%;  Score 15;  DB 2;  Length 114;
  Best Local Similarity   60.0%;  Pred. No. 4.3e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         23 RSLAF 27


RESULT 35
G82537
pterin-4-alpha-carbinolamine dehydratase XF2604 [imported] - Xylella fastidiosa (strai
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82537
```

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82537
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-116 <SIM>
A;Cross-references: UNIPROT:Q9PAB4; UNIPARC:UPI0000131A3A; GB:AE004067; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2604

Query Match 100.0%; Score 15; DB 2; Length 116;
Best Local Similarity 60.0%; Pred. No. 4.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 57 RTLAF 61

RESULT 36
CUPSAM
amicyanin precursor - Methylobacterium extorquens (strain AM1)
C;Species: Methylobacterium extorquens
C;Date: 04-Dec-1986 #sequence_revision 21-Jul-1995 #text_change 05-Oct-2004
C;Accession: A56621; A00295
R;Chistoserdov, A.Y.; Tsygankov, Y.D.; Lidstrom, M.E.
DNA Seq. 2, 53-55, 1991
A;Title: Nucleotide sequence of the amicyanin gene from Methylobacterium extorquens AM1.
A;Reference number: A56621; MUID:92199244; PMID:1802036
A;Accession: A56621
A;Molecule type: DNA
A;Residues: 1-119 <CHI>
A;Cross-references: UNIPROT:P04172; UNIPARC:UPI000012594E; GB:M57963; NID:g150014; PIDN:
A;Note: sequence modified after extraction from NCBI backbone
A;Note: the authors translated the codon CAC for residue 70 as Asn
A;Note: sequence extracted from NCBI backbone (NCBIN:89409, NCBIP:89412)
R;Ambler, R.P.; Tobari, J.
Biochem. J. 232, 451-457, 1985
A;Title: The primary structures of Pseudomonas AM1 amicyanin and pseudoazurin. Two new s
A;Reference number: A90327; MUID:86130354; PMID:4091802
A;Accession: A00295
A;Molecule type: protein
A;Residues: 21-119 <AMB>
A;Cross-references: UNIPARC:UPI0000171E38
C;Comment: This species of Pseudomonas, isolated as an airborne contaminant, uses compou
the true pseudomonads as well as methylotrophs.
C;Superfamily: plastocyanin/Azurin
C;Keywords: copper; electron transfer; metalloprotein; periplasmic space
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-119/Product: amicyanin #status experimental <MAT>
F;67,106,109,112/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 100.0%; Score 15; DB 1; Length 119;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 2 RALAF 6

RESULT 37
E72714
probable ribosomal protein S24 APE1132 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 24-Sep-1999
C;Accession: E72714
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tak
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.,
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72714
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-119 <KAW>
A;Cross-references: UNIPARC:UPI000005DE12; DDBJ:AP000060; NID:g5104188; PIDN:BAA80117
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1132

Query Match 100.0%; Score 15; DB 2; Length 119;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 95 RALSF 99

RESULT 38
G1HUHE
Ig heavy chain V-II region (He) - human
C;Species: Homo sapiens (man)
C;Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 09-Jul-2004
C;Accession: A02093
R;Cunningham, B.A.; Pflumm, M.N.; Rutishauser, U.; Edelman, G.M.
Proc. Natl. Acad. Sci. U.S.A. 64, 997-1003, 1969
A;Title: Subgroups of amino acid sequences in the variable regions of immunoglobulin h
A;Reference number: A02093; MUID:70114712; PMID:5264153
A;Accession: A02093
A;Molecule type: protein
A;Residues: 1-121 <CUN>
A;Cross-references: UNIPROT:P01818; UNIPARC:UPI000012CEEE
C;Comment: This gamma-1 chain was isolated from a myeloma protein.
C;Genetics:
A;Gene: GDB:IGHV@
A;Cross-references: GDB:128528; OMIM:147070
A;Map position: 14q32.33-14q32.33
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: blocked amino end; heterotetramer; immunoglobulin
F;15-100/Domain: immunoglobulin homology <IMM>
F;1/Modified site: blocked amino end (Gln) (probably pyrrolidone carboxylic acid) #sta

Query Match 100.0%; Score 15; DB 1; Length 121;
Best Local Similarity 60.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 104 RTLAF 108

RESULT 39
H82231
hypothetical protein VC1187 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: H82231
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: H82231
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-122 <HEI>
A;Cross-references: UNIPROT:Q9KSR9; UNIPARC:UPI00000C2EDB; GB:AE004198; GB:AE003852; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1187
A;Map position: 1

Query Match 100.0%; Score 15; DB 2; Length 122;
Best Local Similarity 60.0%; Pred. No. 4.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 66 RTLAF 70

RESULT 40
E70547
hypothetical protein Rv0546c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70547
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70547
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-128 <COL>
A;Cross-references: UNIPROT:O06412; UNIPARC:UPI0000031805; GB:Z95558; GB:AL123456; NID:g
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0546c

Query Match 100.0%; Score 15; DB 2; Length 128;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 17 RSLSF 21

RESULT 41
T12924
conserved hypothetical protein bnrdI/yosM - Bacillus subtilis phage SPBc2
C;Species: Bacillus subtilis phage SPBc2
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T12924; G69926
R;Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mauel, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997
A;Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 propha
A;Reference number: Z17583
A;Accession: T12924
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-129 <LAZ>
A;Cross-references: UNIPROT:O31876; UNIPARC:UPI0000060551; EMBL:AF020713; NID:g3025478;
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: G69926
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-129 <KUN>
A;Cross-references: UNIPARC:UPI0000060551; GB:Z99114; GB:AL009126; NID:g2634230; PIDN:C
A;Experimental source: strain 168
C;Genetics: <LA>
A;Gene: bnrdI
C;Genetics: <KU>
A;Gene: yosM
C;Superfamily: Bacillus subtilis conserved hypothetical protein yosM

Query Match 100.0%; Score 15; DB 2; Length 129;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 53 RTLSF 57

RESULT 42
H72627
hypothetical protein APE1480 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: H72627
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-129 <KAW>
A;Cross-references: UNIPROT:Q9YBX0; UNIPARC:UPI000005DF7B; DDBJ:AP000061; NID:g5104821,
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1480
C;Superfamily: Aeropyrum pernix hypothetical protein APE1480

Query Match 100.0%; Score 15; DB 2; Length 129;
Best Local Similarity 60.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 18 RALAF 22

RESULT 43
D48776
polyprotein (E2/NS1 region, HVR1, HVR2) - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: D48776
R;Higashi, Y.; Kakumu, S.; Yoshioka, K.; Wakita, T.; Mizokami, M.; Ohba, K.; Ito, Y.;

Virology 197, 659-668, 1993
A;Title: Dynamics of genome change in the E2/NS1 region of hepatitis C virus in vivo.
A;Reference number: A48776; MUID:94069940; PMID:8249288
A;Accession: D48776
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-133 <HIG>
A;Cross-references: UNIPROT:Q9PXT8; UNIPARC:UPI00000F8865
A;Experimental source: subtype III, patient KS
A;Note: sequence extracted from NCBI backbone (NCBIP:140217)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein

Query Match 100.0%; Score 15; DB 2; Length 133;
Best Local Similarity 60.0%; Pred. No. 5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 13 RTLSF 17

RESULT 44
S15825
agglutinin alpha chain - Osage orange
C;Species: Maclura pomifera (Osage orange)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 31-Dec-2004
C;Accession: S15825
R;Young, N.M.; Johnston, R.A.Z.; Watson, D.C.
FEBS Lett. 282, 382-384, 1991
A;Title: The amino acid sequences of jacalin and the Maclura pomifera agglutinin.
A;Reference number: S15824; MUID:91243835; PMID:2037053
A;Accession: S15825
A;Molecule type: protein
A;Residues: 1-133 <FEB>
A;Cross-references: UNIPROT:P18674; UNIPARC:UPI0000111FE4
C;Superfamily: Mannose-specific lectin
C;Keywords: glycoprotein
F;100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 15; DB 2; Length 133;
Best Local Similarity 60.0%; Pred. No. 5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 82 RSLTF 86

RESULT 45
B30242
stem cell protein ERA-1-399, retinoic acid-induced - mouse
C;Species: Mus musculus (house mouse)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 31-Dec-2004
C;Accession: B30242
R;LaRosa, G.J.; Gudas, L.J.
Mol. Cell. Biol. 8, 3906-3917, 1988
A;Title: Early retinoic acid-induced F9 teratocarcinoma stem cell gene ERA-1: alternate
A;Reference number: A30242; MUID:89127233; PMID:2906112
A;Accession: B30242
A;Molecule type: mRNA
A;Residues: 1-133 <LAR>
A;Cross-references: UNIPROT:P09022; UNIPARC:UPI000002AB6D; GB:M22115; NID:g193047; PIDN:
C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulati

Query Match 100.0%; Score 15; DB 2; Length 133;
Best Local Similarity 60.0%; Pred. No. 5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 117 RSLSF 121

RESULT 46
AF3551
hypothetical protein BMEII0335 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AF3551
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AF3551
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-135 <KUR>
A;Cross-references: UNIPROT:Q8YD42; UNIPARC:UPI000005843E; GB:AE008918; PIDN:AAL53577
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0335
A;Map position: II

Query Match 100.0%; Score 15; DB 2; Length 135;
Best Local Similarity 60.0%; Pred. No. 5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 16 RSLSF 20

RESULT 47
T35211
probable lyase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35211
R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z21572
A;Accession: T35211
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-138 <SEE>
A;Cross-references: UNIPROT:O86701; UNIPARC:UPI00000DADC5; EMBL:AL031515; PIDN:CAA206
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC5C7.04
C;Superfamily: Bacillus probable methylglyoxalase yurT

Query Match 100.0%; Score 15; DB 2; Length 138;
Best Local Similarity 60.0%; Pred. No. 5.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 17 RALAF 21

RESULT 48
A87192
conserved hypothetical protein ML2261 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87192
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroy
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: A87192

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-138 <STO>
A;Cross-references: UNIPROT:Q9CBB6; UNIPARC:UPI00000C6E51; GB:AL450380; NID:g13093763; P
C;Genetics:
A;Gene: ML2261

Query Match 100.0%; Score 15; DB 2; Length 138;
Best Local Similarity 60.0%; Pred. No. 5.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 17 RSLSF 21

RESULT 49
G97080
hypothetical protein CAC1466 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: G97080
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97080
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-139 <KUR>
A;Cross-references: UNIPROT:Q97J20; UNIPARC:UPI00000CA1F2; GB:AE001437; PIDN:AAK79434.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1466

Query Match 100.0%; Score 15; DB 2; Length 139;
Best Local Similarity 60.0%; Pred. No. 5.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 51 RTLAF 55

RESULT 50
H71903
hypothetical protein jhp0657 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: H71903
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: H71903
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-140 <ARN>
A;Cross-references: UNIPROT:Q9ZLC3; UNIPARC:UPI00000D721F; GB:AE001498; GB:AE001439; NID
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0657
C;Superfamily: Helicobacter pylori hypothetical protein jhp0118

Query Match 100.0%; Score 15; DB 2; Length 140;
Best Local Similarity 60.0%; Pred. No. 5.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 3 RSLAF 7

Search completed: May 4, 2006, 06:10:23
Job time : 44 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: May 4, 2006, 06:02:52 ; Search time 228 Seconds
(without alignments)
24.755 Million cell updates/sec

Title: US-09-726-470A-2
Perfect score: 15
Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15	100.0	30	2	Q4YQ95_PLABE	Q4yq95 plasmodium
2	15	100.0	36	2	Q6KGV3_HUMAN	Q6kgv3 homo sapien
3	15	100.0	38	2	Q8JNC7_9ALPH	Q8jnc7 macropodid
4	15	100.0	39	2	Q56W77_ARATH	Q56w77 arabidopsis
5	15	100.0	39	2	Q83JD9_SHIFL	Q83jd9 shigella fl
6	15	100.0	39	2	Q87L29_VIBPA	Q87l29 vibrio para
7	15	100.0	39	2	Q8X417_ECO57	Q8x417 escherichia
8	15	100.0	42	2	Q33005_PINTH	Q33005 pinus thunb
9	15	100.0	46	2	Q8WR26_ANOGA	Q8wr26 anopheles g
10	15	100.0	46	2	Q4T1B2_TETNG	Q4t1b2 tetraodon n
11	15	100.0	47	2	Q5NM51_ZYMMO	Q5nm51 zymomonas m
12	15	100.0	53	2	Q9RAW5_9ACTO	Q9raw5 frankia sp.
13	15	100.0	57	2	Q54G45_DICDI	Q54g45 dictyosteli
14	15	100.0	58	2	Q5JLA2_ORYSA	Q5jla2 oryza sativ
15	15	100.0	60	2	Q7UTY2_RHOBA	Q7uty2 rhodopirell
16	15	100.0	60	2	Q8X8T7_ECO57	Q8x8t7 escherichia
17	15	100.0	63	2	Q71XV1_LISMF	Q71xv1 listeria mo
18	15	100.0	63	2	Q8Y5J7_LISMO	Q8y5j7 listeria mo
19	15	100.0	63	2	Q929V4_LISIN	Q929v4 listeria in
20	15	100.0	64	2	O71115_GVTN	O71115 trichoplusi
21	15	100.0	66	2	Q4X8M0_PLACH	Q4x8m0 plasmodium
22	15	100.0	66	2	Q6KGE3_9CAUD	Q6kge3 bacteriopha
23	15	100.0	68	2	Q8SFT9_HOMGA	Q8sft9 homarus gam
24	15	100.0	68	2	Q4RKC9_TETNG	Q4rkc9 tetraodon n
25	15	100.0	69	2	Q5C2S0_SCHJA	Q5c2s0 schistosoma
26	15	100.0	69	2	Q99IS8_9ZZZZ	Q99is8 uncultured
27	15	100.0	70	2	Q7NSN8_CHRVO	Q7nsn8 chromobacte
28	15	100.0	71	1	Y4UG_RHISN	P55671 rhizobium s
29	15	100.0	71	2	Q62LK3_BURMA	Q62lk3 burkholderi
30	15	100.0	74	2	Q5KRA9_CORGL	Q5kra9 corynebacte
31	15	100.0	75	2	Q5XYJ1_BORGA	Q5xyj1 borrelia ga
32	15	100.0	75	2	Q99IR7_9ZZZZ	Q99ir7 uncultured
33	15	100.0	76	2	Q8MP23_DICDI	Q8mp23 dictyosteli
34	15	100.0	76	2	Q714P1_ENTFC	Q714p1 enterococcu
35	15	100.0	76	2	Q9L8X7_ENTFA	Q9l8x7 enterococcu
36	15	100.0	76	2	Q4QKW7_HAEI8	Q4qkw7 haemophilus
37	15	100.0	77	2	Q54148_SHIFL	Q54148 shigella fl
38	15	100.0	77	2	Q8VSD1_SHIFL	Q8vsd1 shigella fl
39	15	100.0	78	2	Q8XPZ2_RALSO	Q8xpz2 ralstonia s
40	15	100.0	78	2	Q98G90_RHILO	Q98g90 rhizobium l
41	15	100.0	78	2	Q8JTV5_LSDV	Q8jtv5 lumpy skin
42	15	100.0	78	2	Q91MW1_LSDV	Q91mw1 lumpy skin
43	15	100.0	79	1	CATR1_HUMAN	Q13166 homo sapien
44	15	100.0	79	2	Q9I2D4_PSEAE	Q9i2d4 pseudomonas
45	15	100.0	80	2	O14280_SCHPO	O14280 schizosacch
46	15	100.0	82	2	Q6ZH28_ORYSA	Q6zh28 oryza sativ
47	15	100.0	82	2	Q5LFH2_BACFN	Q5lfh2 bacteroides
48	15	100.0	83	2	Q72HW8_THET2	Q72hw8 thermus the
49	15	100.0	83	2	Q6M9S7_PARUW	Q6m9s7 parachlamyd
50	15	100.0	84	2	Q4RRI6_TETNG	Q4rri6 tetraodon n
51	15	100.0	85	2	Q7R3K7_GIALA	Q7r3k7 giardia lam
52	15	100.0	85	2	Q89B55_BRAJA	Q89b55 bradyrhizob
53	15	100.0	85	2	Q4T498_TETNG	Q4t498 tetraodon n
54	15	100.0	86	2	Q7RXB9_NEUCR	Q7rxb9 neurospora
55	15	100.0	86	2	Q53CV7_9GAMA	Q53cv7 macaca fusc
56	15	100.0	87	2	Q5YTL3_NOCFA	Q5ytl3 nocardia fa
57	15	100.0	87	2	Q6D5X8_ERWCT	Q6d5x8 erwinia car
58	15	100.0	87	2	Q85UG3_9PERC	Q85ug3 sebastolobu
59	15	100.0	88	2	Q6PJL9_HUMAN	Q6pjl9 homo sapien
60	15	100.0	90	2	Q8M241_PRANA	Q8m241 praomys nat
61	15	100.0	90	2	Q8LW23_MASCO	Q8lw23 mastomys co
62	15	100.0	90	2	Q85345_9POXV	Q85345 vaccinia vi
63	15	100.0	91	2	Q5Z906_ORYSA	Q5z906 oryza sativ
64	15	100.0	91	2	Q6YYV6_ORYSA	Q6yyv6 oryza sativ
65	15	100.0	91	2	Q5NN06_ZYMMO	Q5nn06 zymomonas m
66	15	100.0	91	2	Q64MG1_BACFR	Q64mg1 bacteroides
67	15	100.0	91	2	Q8M242_MASCO	Q8m242 mastomys co
68	15	100.0	91	2	Q8M239_PRANA	Q8m239 praomys nat
69	15	100.0	92	2	Q56BC4_9CAUD	Q56bc4 enterobacte
70	15	100.0	92	2	Q56ZY0_ARATH	Q56zy0 arabidopsis
71	15	100.0	93	2	Q4ZT23_PSESY	Q4zt23 pseudomonas
72	15	100.0	93	2	Q5PM22_SALPA	Q5pm22 salmonella
73	15	100.0	93	2	Q8Z8C3_SALTI	Q8z8c3 salmonella
74	15	100.0	93	2	Q8ZQT8_SALTY	Q8zqt8 salmonella
75	15	100.0	93	2	Q8M240_PRANA	Q8m240 praomys nat
76	15	100.0	95	2	O32936_MYCLE	O32936 mycobacteri
77	15	100.0	95	2	Q7MBN9_VIBVY	Q7mbn9 vibrio vuln
78	15	100.0	95	2	Q7MZK3_PHOLL	Q7mzk3 photorhabdu
79	15	100.0	95	2	Q9CM66_PASMU	Q9cm66 pasteurella
80	15	100.0	97	1	YBGE_ECOLI	P37343 escherichia
81	15	100.0	97	2	Q8TRU8_METAC	Q8tru8 methanosarc
82	15	100.0	97	2	Q57RK8_SALCH	Q57rk8 salmonella
83	15	100.0	97	2	Q65UN5_MANSM	Q65un5 mannheimia
84	15	100.0	98	2	Q92ZE9_RHIME	Q92ze9 rhizobium m
85	15	100.0	98	2	Q7VBD9_PROMA	Q7vbd9 prochloroco
86	15	100.0	98	2	Q981R5_RHILO	Q981r5 rhizobium l
87	15	100.0	99	2	Q8J2X7_PYRAB	Q8j2x7 pyrococcus
88	15	100.0	100	1	YJCE_BACSU	O31627 bacillus su
89	15	100.0	100	2	Q4SZM0_TETNG	Q4szm0 tetraodon n
90	15	100.0	101	2	Q7LIF1_YEAST	Q7lif1 saccharomyc
91	15	100.0	101	2	Q7WXG2_ALCEU	Q7wxg2 alcaligenes
92	15	100.0	101	2	Q6LP42_PHOPR	Q6lp42 photobacter
93	15	100.0	101	2	Q73LU7_TREDE	Q73lu7 treponema d
94	15	100.0	101	2	Q928A0_LISIN	Q928a0 listeria in
95	15	100.0	101	2	Q8Y4F1_LISMO	Q8y4f1 listeria mo
96	15	100.0	101	2	Q71WT5_LISMF	Q71wt5 listeria mo
97	15	100.0	102	1	RS24_AERPE	Q9ycy0 aeropyrum p
98	15	100.0	102	2	Q93KR2_YEREN	Q93kr2 yersinia en
99	15	100.0	102	2	Q5SJY5_THET8	Q5sjy5 thermus the
100	15	100.0	102	2	Q5F4C9_CHICK	Q5f4c9 gallus gall

ALIGNMENTS

```
RESULT 1
Q4YQ95_PLABE
ID   Q4YQ95_PLABE PRELIMINARY;      PRT;    30 AA.
AC   Q4YQ95;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PB107776.00.0;
OS   Plasmodium berghei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5821;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAI01003005; CAH99814.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   30 AA;  3718 MW;  A608D7A02BCC490F CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 7.5e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          8 RALSF 12


RESULT 2
Q6KGV3_HUMAN
ID   Q6KGV3_HUMAN PRELIMINARY;      PRT;    36 AA.
AC   Q6KGV3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome P450 variant 3A7 (Fragment).
GN   Name=CYP3A7;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20578884; PubMed=11137287; DOI=10.1016/S0378-1119(00)00470-4;
RA   Finta C., Zaphiropoulos P.G.;
RT   "The human cytochrome P450 3A locus. Gene evolution by capture of
RT   downstream exons.";
RL   Gene 260:13-23(2000).
DR   EMBL; AF315322; AAG48617.1; -; Genomic_DNA.
DR   EMBL; AF315321; AAG48617.1; JOINED; Genomic_DNA.
DR   Ensembl; ENSG00000160870; Homo sapiens.
FT   NON_TER       1       1
SQ   SEQUENCE   36 AA;  4354 MW;  32798EF0F1646523 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 36;
  Best Local Similarity   60.0%;  Pred. No. 8.9e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         13 RSLAF 17


RESULT 3
Q8JNC7_9ALPH
ID   Q8JNC7_9ALPH PRELIMINARY;      PRT;    38 AA.
AC   Q8JNC7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Protein kinase (Fragment).
GN   Name=US3;
OS   Macropodid herpesvirus 2 (dorcopsis wallaby herpesvirus).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=83440;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Thomson D., Smith G.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY048540; AAL13141.1; -; Genomic_DNA.
DR   GO; GO:0016301; F:kinase activity; IEA.
KW   Kinase.
FT   NON_TER       1       1
SQ   SEQUENCE   38 AA;  4596 MW;  C6708AA1DDCFE4AD CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 38;
  Best Local Similarity   60.0%;  Pred. No. 9.3e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         15 RALTF 19


RESULT 4
Q56W77_ARATH
ID   Q56W77_ARATH PRELIMINARY;      PRT;    39 AA.
AC   Q56W77;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
RA   Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
RA   Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
RA   Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RT   "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK222170; BAD95288.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   39 AA;  4740 MW;  5B2A276901EC0D8D CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 39;
  Best Local Similarity   60.0%;  Pred. No. 9.6e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         32 RSLSF 36


RESULT 5
Q83JD9_SHIFL
```

```
ID   Q83JD9_SHIFL PRELIMINARY;      PRT;    39 AA.
AC   Q83JD9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=SF3292;
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
DR   EMBL; AE005674; AAN44756.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   39 AA;  4694 MW;  F56DF3E12C2D5A19 CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 39;
  Best Local Similarity   60.0%;  Pred. No. 9.6e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         28 RALAF 32


RESULT 6
Q87L29_VIBPA
ID   Q87L29_VIBPA PRELIMINARY;      PRT;    39 AA.
AC   Q87L29;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein VP2787.
GN   OrderedLocusNames=VP2787;
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=RIMD 2210633 / Serotype O3:K6;
RX   MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT   distinct from that of V. cholerae.";
RL   Lancet 361:743-749(2003).
DR   EMBL; BA000031; BAC61050.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   39 AA;  4483 MW;  CC3BC98266A617FE CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 39;
  Best Local Similarity   60.0%;  Pred. No. 9.6e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         26 RSLSF 30


RESULT 7
Q8X417_ECO57
ID   Q8X417_ECO57 PRELIMINARY;      PRT;    39 AA.
AC   Q8X417;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   , complete genome.
GN   OrderedLocusNames=z4614;
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX   MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
DR   EMBL; AE005174; AAG58382.1; -; Genomic_DNA.
DR   PIR; B85990; B85990.
KW   Complete proteome.
SQ   SEQUENCE   39 AA;  4604 MW;  9607DF8C26905A1D CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 39;
  Best Local Similarity   60.0%;  Pred. No. 9.6e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         28 RALAF 32


RESULT 8
Q33005_PINTH
ID   Q33005_PINTH PRELIMINARY;      PRT;    42 AA.
AC   Q33005;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   ORF42g.
OS   Pinus thunbergii (Green pine) (Japanese black pine).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
OX   NCBI_TaxID=3350;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=92212283; PubMed=1557027;
RA   Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,
RA   Wakasugi T., Sugiura M.;
RT   "Chloroplast DNA of black pine retains a residual inverted repeat
RT   lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and
RT   trnH and the absence of rps16.";
RL   Mol. Gen. Genet. 232:206-214(1992).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=95094312; PubMed=8001170; DOI=10.1007/BF00313804;
RA   Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugiura M.;
RT   "A new gene encoding tRNA pro (GGG) is present in the chloroplast
RT   genome of black pine: a compilation of 32 tRNA genes from black pine
RT   chloroplasts.";
RL   Curr. Genet. 26:153-158(1994).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=95024047; PubMed=7937893;
RA   Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
RA   Sugiura M.;
```

```
RT   "Loss of all ndh genes as determined by sequencing the entire
RT   chloroplast genome of the black pine Pinus thunbergii.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
DR   EMBL; D17510; BAA04457.1; -; Genomic_DNA.
DR   PIR; T07581; T07581.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
SQ   SEQUENCE   42 AA;  4945 MW;  932D81DDA0604964 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 42;
  Best Local Similarity   60.0%;  Pred. No. 1e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         14 RSLSF 18


RESULT 9
Q8WR26_ANOGA
ID   Q8WR26_ANOGA PRELIMINARY;      PRT;    46 AA.
AC   Q8WR26;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Anopheles gambiae (African malaria mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC   Anophelinae; Anopheles.
OX   NCBI_TaxID=7165;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22120133; PubMed=12124367;
RA   Francischetti I.M., Valenzuela J.G., Pham V.M., Garfield M.K.,
RA   Ribeiro J.M.;
RT   "Toward a catalog for the transcripts and proteins (sialome) from the
RT   salivary gland of the malaria vector Anopheles gambiae.";
RL   J. Exp. Biol. 205:2429-2451(2002).
DR   EMBL; AF457561; AAL68791.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   46 AA;  5266 MW;  9D378CAAF35E4CD2 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 46;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         16 RALSF 20


RESULT 10
Q4T1B2_TETNG
ID   Q4T1B2_TETNG PRELIMINARY;      PRT;    46 AA.
AC   Q4T1B2;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome undetermined SCAF10686, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00008913001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01010686; CAF93320.1; -; Genomic_DNA.
FT   NON_TER      46     46
SQ   SEQUENCE   46 AA;  5314 MW;  8E65040019D2CE69 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 46;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         21 RALTF 25


RESULT 11
Q5NM51_ZYMMO
ID   Q5NM51_ZYMMO PRELIMINARY;      PRT;    47 AA.
AC   Q5NM51;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=ZMO1585;
OS   Zymomonas mobilis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Zymomonas.
OX   NCBI_TaxID=542;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=ATCC 31821 / ZM4 / CP4;
RX   PubMed=15592456; DOI=10.1038/nbt1045;
RA   Seo J.-S., Chong H., Park H.S., Yoon K.-O., Jung C., Kim J.J.,
RA   Hong J.H., Kim H., Kim J.-H., Kil J.-I., Park C.J., Oh H.-M.,
RA   Lee J.-S., Jin S.-J., Um H.-W., Lee H.-J., Oh S.-J., Kim J.Y.,
RA   Kang H.L., Lee S.Y., Lee K.J., Kang H.S.;
RT   "The genome sequence of the ethanologenic bacterium Zymomonas mobilis
RT   ZM4.";
RL   Nat. Biotechnol. 23:63-68(2005).
DR   EMBL; AE008692; AAV90209.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   47 AA;  5367 MW;  B3B30EF2D3EF0395 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 47;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          6 RSLAF 10


RESULT 12
Q9RAW5_9ACTO
ID   Q9RAW5_9ACTO PRELIMINARY;      PRT;    53 AA.
AC   Q9RAW5;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative dienelactone hydrolase-like protein (Fragment).
OS   Frankia sp. ACN14a/ts-r.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Frankineae; Frankiaceae; Frankia.
OX   NCBI_TaxID=92643;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ACN14a/ts-r;
RX   MEDLINE=20117996; PubMed=10652097;
RA   Pouch M.N., Cournoyer B., Baumeister W.;
RT   "Characterization of the 20S proteasome from the actinomycete
RT   Frankia.";
RL   Mol. Microbiol. 35:368-377(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ACN14a/ts-r;
RA   Pouch M.-N., Cournoyer B., Baumeister W.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF142435; AAF14269.1; -; Genomic_DNA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR002925; Dienelactn_hydro.
DR   Pfam; PF01738; DLH; 1.
KW   Hydrolase.
FT   NON_TER       1       1
SQ   SEQUENCE   53 AA;  5748 MW;  CAF445055B8B1E87 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 53;
  Best Local Similarity   60.0%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         41 RALAF 45


RESULT 13
Q54G45_DICDI
ID   Q54G45_DICDI PRELIMINARY;      PRT;    57 AA.
AC   Q54G45;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=DDB0188876;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAFI01000234; EAL62222.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   57 AA;  6817 MW;  C59D67947350CC13 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 57;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         46 RSLTF 50


RESULT 14
Q5JLA2_ORYSA
ID   Q5JLA2_ORYSA PRELIMINARY;      PRT;    58 AA.
AC   Q5JLA2;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein P0480C01.16.
GN   Name=P0480C01.16;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA   Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA   Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA   Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA   Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA   Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA   Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA   Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA   Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA   Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA   Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA   Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA   Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA   Yano M., Jiang J., Gojobori T.;
RT   "The genome sequence and structure of rice chromosome 1.";
RL   Nature 420:312-316(2002).
DR   EMBL; AP003453; BAD87755.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   58 AA;  6494 MW;  6F1957C17710F1A3 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 58;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          7 RALSF 11


RESULT 15
Q7UTY2_RHOBA
ID   Q7UTY2_RHOBA PRELIMINARY;      PRT;    60 AA.
AC   Q7UTY2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=RB3623;
OS   Rhodopirellula baltica.
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC   Planctomycetaceae; Pirellula.
```

```
OX   NCBI_TaxID=117;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=1;
RX   MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA   Schlesner H., Amann R., Reinhardt R.;
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294139; CAD73302.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   60 AA;  6661 MW;  22B188C360C9708F CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 60;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         56 RSLTF 60


RESULT 16
Q8X8T7_ECO57
ID   Q8X8T7_ECO57 PRELIMINARY;      PRT;    60 AA.
AC   Q8X8T7; Q7ACV4;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein ECs2740.
GN   OrderedLocusNames=ECs2740, z2370;
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX   MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
DR   EMBL; AE005174; AAG56416.1; -; Genomic_DNA.
DR   EMBL; BA000007; BAB36163.1; -; Genomic_DNA.
DR   PIR; D85744; D85744.
DR   PIR; D90971; D90971.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   60 AA;  6544 MW;  AE13AA97B7255109 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 60;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         29 RALAF 33


RESULT 17
Q71XV1_LISMF
ID   Q71XV1_LISMF PRELIMINARY;      PRT;    63 AA.
AC   Q71XV1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=LMOf2365_2094;
OS   Listeria monocytogenes (serotype 4b / strain F2365).
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=265669;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15115801; DOI=10.1093/nar/gkh562;
RA   Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA   Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
RA   Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,
RA   Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA   Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,
RA   Forberger H.A., Tran B., Kathariou S., Wonderling L.D., Uhlich G.A.,
RA   Bayles D.O., Luchansky J.B., Fraser C.M.;
RT   "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
RT   borne pathogen Listeria monocytogenes reveal new insights into the
RT   core genome components of this species.";
RL   Nucleic Acids Res. 32:2386-2395(2004).
DR   EMBL; AE017329; AAT04864.1; -; Genomic_DNA.
DR   TIGR; LMOf2365_2094; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   63 AA;  7504 MW;  A31E13B6C0BD3050 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 63;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         39 RTLTF 43


RESULT 18
Q8Y5J7_LISMO
ID   Q8Y5J7_LISMO PRELIMINARY;      PRT;    63 AA.
AC   Q8Y5J7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Lmo2063 protein.
GN   OrderedLocusNames=lmo2063;
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=EGD-e / Serovar 1/2a;
RX   MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL591982; CAD00141.1; -; Genomic_DNA.
```

```
DR   PIR; AG1332; AG1332.
DR   ListiList; LMO2063; -.
KW   Complete proteome.
SQ   SEQUENCE   63 AA;  7504 MW;  A31E07098B5D3050 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 63;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         39 RTLTF 43


RESULT 19
Q929V4_LISIN
ID   Q929V4_LISIN PRELIMINARY;      PRT;    63 AA.
AC   Q929V4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Lin2169 protein.
GN   OrderedLocusNames=lin2169;
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CLIP 11262 / Serovar 6a;
RX   MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596171; CAC97399.1; -; Genomic_DNA.
DR   PIR; AG1703; AG1703.
DR   ListiList; LIN2169; -.
KW   Complete proteome.
SQ   SEQUENCE   63 AA;  7418 MW;  B7A15C4291E86115 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 63;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         39 RTLTF 43


RESULT 20
O71115_GVTN
ID   O71115_GVTN PRELIMINARY;      PRT;    64 AA.
AC   O71115;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   P7.41.
GN   Name=ORF507;
OS   Trichoplusia ni granulosis virus (TnGV) (Trichoplusia ni
OS   granulovirus).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX   NCBI_TaxID=10462;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=98264509; PubMed=9603347;
RA   Bideshi D.K., Hice R.H., Ge B., Federici B.A.;
RT   "Molecular characterization and expression of the Trichoplusia ni
RT   granulovirus helicase gene.";
RL   J. Gen. Virol. 79:1309-1319(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Bideshi D.B., Hice R.H., Ge B., Federici B.A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF032994; AAC40853.1; -; Genomic_DNA.
SQ   SEQUENCE   64 AA;  7416 MW;  AA541DB3DDC74ED6 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 64;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         15 RALSF 19


RESULT 21
Q4X8M0_PLACH
ID   Q4X8M0_PLACH PRELIMINARY;      PRT;    66 AA.
AC   Q4X8M0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC404933.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01008836; CAH86756.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
FT   NON_TER      66     66
SQ   SEQUENCE   66 AA;  7233 MW;  540CB6FEA7CEC7E2 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 66;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          1 RALAF 5


RESULT 22
Q6KGE3_9CAUD
ID   Q6KGE3_9CAUD PRELIMINARY;      PRT;    66 AA.
AC   Q6KGE3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Bacteriophage Felix 01.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales.
OX   NCBI_TaxID=77775;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sriranganathan N., Whichard J.M., Pierson F.W., Kapur V., Weigt L.A.;
RT   "Bacteriophage Felix O1: Genetic Characterization.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF320576; AAQ14689.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   66 AA;  7831 MW;  C0D3134561842994 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 66;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          3 RSLSF 7


RESULT 23
Q8SFT9_HOMGA
ID   Q8SFT9_HOMGA PRELIMINARY;      PRT;    68 AA.
AC   Q8SFT9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Cytochrome b (Fragment).
GN   Name=CYTB;
OS   Homarus gammarus (European lobster) (Homarus vulgaris).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC   Nephropoidea; Nephropidae; Homarus.
OX   NCBI_TaxID=6707;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=14730430; DOI=10.1007/s10126-002-0097-7;
RA   Katsares V., Apostolidis A., Triantafyllidis A., Kouvatsi A.,
RA   Triantaphyllidis C.;
RT   "Development of mitochondrial DNA primers for use with homarid
RT   lobsters.";
RL   Mar. Biotechnol. 5:469-479(2003).
CC   -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC       cytochrome c1 and the Rieske protein (By similarity).
DR   EMBL; AF494203; AAM15924.1; -; Genomic_DNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR   GO; GO:0005739; C:mitochondrion; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR005798; Cytb_b6_C.
DR   Pfam; PF00032; Cytochrom_B_C; 1.
DR   PROSITE; PS51003; CYTB_CTER; 1.
KW   Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW   Respiratory chain; Transmembrane; Transport.
FT   NON_TER       1      1
SQ   SEQUENCE   68 AA;  8074 MW;  99EFF029E09DC1D0 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 68;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          3 RSLTF 7


RESULT 24
Q4RKC9_TETNG
ID   Q4RKC9_TETNG PRELIMINARY;      PRT;    68 AA.
AC   Q4RKC9;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 21 SCAF15029, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00033014001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01015029; CAG11153.1; -; Genomic_DNA.
FT   NON_TER       68     68
SQ   SEQUENCE   68 AA;  7223 MW;  3AFE0CF0917521E6 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 68;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         39 RTLAF 43


RESULT 25
Q5C2S0_SCHJA
ID   Q5C2S0_SCHJA PRELIMINARY;      PRT;    69 AA.
AC   Q5C2S0;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Schistosoma japonicum (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6182;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Han Z.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY810166; AAX26055.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   69 AA;  8015 MW;  7B2222568B755911 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 69;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          4 RXLXF 8
              | | |
Db         58 RTLSF 62


RESULT 26
Q99IS8_9ZZZZ
ID   Q99IS8_9ZZZZ PRELIMINARY;      PRT;    69 AA.
AC   Q99IS8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   uncultured organism.
OC   unclassified; environmental samples.
OX   NCBI_TaxID=155900;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Stokes H.W., Nield B.S., Mabbutt B.C., Nevalainen H., Holmes A.J.,
RA   Gillings M.R.;
RT   "Novel and diverse integron-like gene cassettes are prevalent in
RT   natural environments.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF349105; AAK28612.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   69 AA;  7736 MW;  EB73269523C2B349 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 69;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         21 RALSF 25


RESULT 27
Q7NSN8_CHRVO
ID   Q7NSN8_CHRVO PRELIMINARY;      PRT;    70 AA.
AC   Q7NSN8;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=CV3382;
OS   Chromobacterium violaceum.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Chromobacterium.
OX   NCBI_TaxID=536;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 12472 / DSM 30191;
RX   MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RA   Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA   Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA   Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.F.,
RA   Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA   Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA   Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,
RA   Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA   Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
RA   Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA   Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA   Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA   Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA   Grattapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA   Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA   Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA   di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
RA   Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA   Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA   Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA   Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA   Santos E.B.P., Santos F.R., Schneider M.P.C., Seuanez H.N.,
RA   Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA   Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA   Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA   Vettore A., Wassem R., Zaha A., Simpson A.J.G.;
RT   "The complete genome sequence of Chromobacterium violaceum reveals
RT   remarkable and exploitable bacterial adaptability.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR   EMBL; AE016922; AAQ61046.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   70 AA;  7585 MW;  1127465F5AF93628 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 70;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         55 RALSF 59


RESULT 28
Y4UG_RHISN
ID   Y4UG_RHISN     STANDARD;      PRT;    71 AA.
AC   P55671;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Hypothetical 7.8 kDa protein y4uG.
GN   ORFNames=y4uG;
OS   Rhizobium sp. (strain NGR234).
OG   Plasmid sym pNGR234a.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=394;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=97305956; PubMed=9163424;
RA   Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA   Perret X.;
RT   "Molecular basis of symbiosis between Rhizobium and legumes.";
RL   Nature 387:394-401(1997).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=96389014; PubMed=8796346;
RA   Freiberg C., Perret X., Broughton W.J., Rosenthal A.;
RT   "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.
RT   NGR234 using dye terminators and a thermostable 'sequenase': a
RT   beginning.";
RL   Genome Res. 6:590-600(1996).
CC   ---------------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   EMBL; Z68203; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR   EMBL; AE000099; AAB91879.1; -; Genomic_DNA.
KW   Hypothetical protein; Plasmid; Transmembrane.
FT   TRANSMEM     12     32       Potential.
SQ   SEQUENCE   71 AA;  7769 MW;  655F2FDA41049001 CRC64;

  Query Match             100.0%;  Score 15;  DB 1;  Length 71;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          8 RSLSF 12


RESULT 29
```

```
Q62LK3_BURMA
ID   Q62LK3_BURMA PRELIMINARY;      PRT;    71 AA.
AC   Q62LK3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=BMA0634;
OS   Burkholderia mallei (Pseudomonas mallei).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Burkholderia.
OX   NCBI_TaxID=13373;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 23344;
RX   PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA   Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA   Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA   Daugherty S.C., Davidsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA   Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA   Madupu R., Mohammoud Y., Nelson W.C., Radune D., Romero C.M.,
RA   Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA   Zafar N., Zhou L., Fraser C.M.;
RT   "Structural flexibility in the Burkholderia mallei genome.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR   EMBL; CP000010; AAU49688.1; -; Genomic_DNA.
DR   TIGR; BMA0634; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   71 AA;  7537 MW;  D253CF3D92034AD4 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 71;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         11 RALTF 15


RESULT 30
Q5KRA9_CORGL
ID   Q5KRA9_CORGL PRELIMINARY;      PRT;    74 AA.
AC   Q5KRA9;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Putative ribosomal pseudouridine synthase (Fragment).
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=R;
RA   Suzuki N., Okayama S., Nonaka H., Tsuge Y., Inui M., Yukawa H.;
RT   "Large-Scale Engineering of the Corynebacterium glutamicum Genome.";
RL   Appl. Environ. Microbiol. 71:3369-3372(2005).
DR   EMBL; AB193038; BAD84160.1; -; Genomic_DNA.
DR   GO; GO:0009982; F:pseudouridine synthase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0006396; P:RNA processing; IEA.
DR   InterPro; IPR006145; PseudoU_synth.
DR   Pfam; PF00849; PseudoU_synth_2; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   74 AA;  8349 MW;  D04A8CE5DB8E9A86 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 74;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         50 RTLTF 54


RESULT 31
Q5XYJ1_BORGA
ID   Q5XYJ1_BORGA PRELIMINARY;      PRT;    75 AA.
AC   Q5XYJ1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=BGP290;
OS   Borrelia garinii PBi.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC   Borrelia burgdorferi group.
OX   NCBI_TaxID=290434;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PBi;
RA   Glockner G., Lehmann R., Romualdi A., Pradella S.,
RA   Schulte-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzer M.;
RT   "Comparative analysis of the Borrelia garinii genome.";
RL   Nucleic Acids Res. 32:6038-6046(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PBi;
RA   Gloeckner G., Schilhabel M., Lehmann R., Platzer M.;
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY722938; AAU86141.1; -; Genomic_DNA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003678; F:DNA helicase activity; IEA.
DR   GO; GO:0006260; P:DNA replication; IEA.
DR   InterPro; IPR007694; DnaB_C.
DR   ProDom; PD332834; DnaB_C; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   75 AA;  8609 MW;  92784FAB78DF9451 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 75;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         47 RALAF 51


RESULT 32
Q99IR7_9ZZZZ
ID   Q99IR7_9ZZZZ PRELIMINARY;      PRT;    75 AA.
AC   Q99IR7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   uncultured organism.
OC   unclassified; environmental samples.
OX   NCBI_TaxID=155900;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Stokes H.W., Nield B.S., Mabbutt B.C., Nevalainen H., Holmes A.J.,
RA   Gillings M.R.;
RT   "Novel and diverse integron-like gene cassettes are prevalent in
RT   natural environments.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF349110; AAK28623.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   75 AA;  8285 MW;  B0DAC6D7E3CF39C2 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 75;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
```

```
Db         27 RALSF 31


RESULT 33
Q8MP23_DICDI
ID   Q8MP23_DICDI PRELIMINARY;      PRT;    76 AA.
AC   Q8MP23; Q8T172;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=DDB0167780;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA   Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA   Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RX   MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA   Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA   Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA   Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL   Nature 418:79-85(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Baumgart C.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
DR   EMBL; AC123513; AAM44374.1; -; Genomic_DNA.
DR   EMBL; AC117070; AAM09344.2; -; Genomic_DNA.
DR   EMBL; AAFI01000027; EAL70167.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   76 AA;  8901 MW;  A4A875C0EC898A20 CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 76;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         67 RTLAF 71


RESULT 34
Q714P1_ENTFC
ID   Q714P1_ENTFC PRELIMINARY;      PRT;    76 AA.
AC   Q714P1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Enterococcus faecium (Streptococcus faecium).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1352;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=BM4524;
RX   MEDLINE=22983133; PubMed=14617162;
RX   DOI=10.1046/j.1365-2958.2003.03771.x;
RA   Depardieu F., Courvalin P., Msadek T.;
RT   "A six amino acid deletion, partially overlapping the VanSB G2 ATP-
RT   binding motif, leads to constitutive glycopeptide resistance in VanB-
RT   type Enterococcus faecium.";
RL   Mol. Microbiol. 50:1069-1083(2003).
DR   EMBL; AF550667; AAQ12899.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   76 AA;  8698 MW;  C628291E1C80D050 CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 76;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          7 RALSF 11


RESULT 35
Q9L8X7_ENTFA
ID   Q9L8X7_ENTFA PRELIMINARY;      PRT;    76 AA.
AC   Q9L8X7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Enterococcus faecalis (Streptococcus faecalis).
OG   Plasmid pIP834.
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=BM4382;
RX   MEDLINE=20307505; PubMed=10846226;
RA   Garnier F., Taourit S., Glaser P., Courvalin P., Galimand M.;
RT   "Characterization of transposon Tn1549, conferring VanB-type
RT   resistance in Enterococcus spp.";
RL   Microbiology 146:1481-1489(2000).
DR   EMBL; AF192329; AAF72366.1; -; Genomic_DNA.
KW   Hypothetical protein; Plasmid.
SQ   SEQUENCE   76 AA;  8698 MW;  C628291E1C80D050 CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 76;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          7 RALSF 11


RESULT 36
Q4QKW7_HAEI8
ID   Q4QKW7_HAEI8 PRELIMINARY;      PRT;    76 AA.
AC   Q4QKW7;
```

```
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=NTHI1521;
OS   Haemophilus influenzae (strain 86-028NP).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=281310;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA   Harrison A., Dyer D.W., Gillaspy A., Ray W.C., Mungur R., Carson M.B.,
RA   Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,
RA   Munson R.S. Jr.;
RT   "Genomic sequence of an otitis media isolate of nontypeable
RT   Haemophilus influenzae: comparative study with H. influenzae serotype
RT   d, strain KW20.";
RL   J. Bacteriol. 187:4627-4636(2005).
DR   EMBL; CP000057; AAX88330.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   76 AA;  8359 MW;  DA2EBA01A768CF88 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 76;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         15 RTLAF 19


RESULT 37
Q54148_SHIFL
ID   Q54148_SHIFL PRELIMINARY;      PRT;    77 AA.
AC   Q54148;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Invasion plasmid pWR100 ipgH ORFA and ipgH ORFB genes, complete cds
DE   (Orf, hypothetical).
GN   Name=S0089;
OS   Shigella flexneri.
OG   Plasmid invasion plasmid pWR100, and Plasmid virulence plasmid pWR501.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=M90T-W; PLASMID=invasion plasmid pWR100;
RX   MEDLINE=92167809; PubMed=1791758;
RA   Venkatesan M.M., Buysse J.M., Hartman A.B.;
RT   "Sequence variation in two ipaH genes of Shigella flexneri 5 and
RT   homology to the LRG-like family of proteins.";
RL   Mol. Microbiol. 5:2435-2445(1991).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=M90T-W; PLASMID=invasion plasmid pWR100;
RX   MEDLINE=97074644; PubMed=8917071; DOI=10.1016/0378-1119(96)00115-1;
RA   Venkatesan M.M., Alexander W.A., Fernandez-Prada C.;
RT   "A Shigella flexneri invasion plasmid gene, ipgH, with homology to
RT   IS629 and sequences encoding bacterial sugar phosphate transport
RT   proteins.";
RL   Gene 175:23-27(1996).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   PLASMID=virulence plasmid pWR501;
RX   MEDLINE=21189246; PubMed=11292750;
RX   DOI=10.1128/IAI.69.5.3271-3285.2001;
RA   Venkatesan M.M., Goldberg M.B., Rose D.J., Grotbeck E.J., Burland V.,
RA   Blattner F.R.;
RT   "Complete DNA sequence and analysis of the large virulence plasmid of
RT   Shigella flexneri.";
RL   Infect. Immun. 69:3271-3285(2001).
DR   EMBL; U28354; AAC44576.1; -; Genomic_DNA.
DR   EMBL; AF348706; AAK18399.1; -; Genomic_DNA.
DR   PIR; JC5052; JC5052.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   InterPro; IPR002611; IstB_ATPbind.
DR   Pfam; PF01695; IstB; 1.
KW   Hypothetical protein; Plasmid.
SQ   SEQUENCE   77 AA;  8875 MW;  01C1DFDA949974C3 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 77;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         49 RSLSF 53


RESULT 38
Q8VSD1_SHIFL
ID   Q8VSD1_SHIFL PRELIMINARY;      PRT;    77 AA.
AC   Q8VSD1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   IS100 ORF2.
GN   OrderedLocusNames=CP0214;
OS   Shigella flexneri.
OG   Plasmid pCP301.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=301;
RX   MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
DR   EMBL; AF386526; AAL72411.1; -; Genomic_DNA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   InterPro; IPR002611; IstB_ATPbind.
DR   Pfam; PF01695; IstB; 1.
KW   Complete proteome; Plasmid.
SQ   SEQUENCE   77 AA;  8860 MW;  01D2426C5C696EEA CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 77;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         49 RSLSF 53


RESULT 39
Q8XPZ2_RALSO
ID   Q8XPZ2_RALSO PRELIMINARY;      PRT;    78 AA.
AC   Q8XPZ2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein RSp1494.
GN   OrderedLocusNames=RSp1494; ORFNames=RS03050;
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OG   Plasmid megaplasmid.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
```

```
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646085; CAD18645.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein; Plasmid.
SQ   SEQUENCE   78 AA;  8533 MW;  7D73171150B8B7B5 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 78;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         12 RSLAF 16


RESULT 40
Q98G90_RHILO
ID   Q98G90_RHILO PRELIMINARY;      PRT;    78 AA.
AC   Q98G90;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Msl3438 protein.
GN   OrderedLocusNames=msl3438;
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; BA000012; BAB50326.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   78 AA;  8168 MW;  247DD71269947291 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 78;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         55 RALSF 59


RESULT 41
Q8JTV5_LSDV
ID   Q8JTV5_LSDV PRELIMINARY;      PRT;    78 AA.
AC   Q8JTV5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein LW046.
GN   Name=LW046;
OS   Lumpy skin disease virus (LSDV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Capripoxvirus.
OX   NCBI_TaxID=59509;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22710395; PubMed=12827464;
RA   Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Abolnik C., Lu Z.,
RA   Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;
RT   "Comparative sequence analysis of the South African vaccine strain and
RT   two virulent field isolates of Lumpy skin disease virus.";
RL   Arch. Virol. 148:1335-1356(2003).
DR   EMBL; AF409138; AAN02771.1; -; Genomic_DNA.
DR   InterPro; IPR006803; Pox_I5.
DR   Pfam; PF04713; Pox_I5; 1.
DR   PIRSF; PIRSF003768; VAC_I5L; 1.
DR   ProDom; PD012198; Pox_I5; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   78 AA;  8797 MW;  4E991CFF366D4D56 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 78;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         42 RSLTF 46


RESULT 42
Q91MW1_LSDV
ID   Q91MW1_LSDV PRELIMINARY;      PRT;    78 AA.
AC   Q91MW1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein LSDV046 (Hypothetical protein LD046).
GN   Name=LSDV046; Synonyms=LD046;
OS   Lumpy skin disease virus (LSDV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Capripoxvirus.
OX   NCBI_TaxID=59509;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21329495; PubMed=11435593;
RX   DOI=10.1128/JVI.75.15.7122-7130.2001;
RA   Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "Genome of lumpy skin disease virus.";
RL   J. Virol. 75:7122-7130(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22710395; PubMed=12827464;
RA   Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Abolnik C., Lu Z.,
RA   Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;
RT   "Comparative sequence analysis of the South African vaccine strain and
RT   two virulent field isolates of Lumpy skin disease virus.";
RL   Arch. Virol. 148:1335-1356(2003).
DR   EMBL; AF325528; AAK85007.1; -; Genomic_DNA.
DR   EMBL; AF409137; AAN02614.1; -; Genomic_DNA.
DR   InterPro; IPR006803; Pox_I5.
DR   Pfam; PF04713; Pox_I5; 1.
DR   PIRSF; PIRSF003768; VAC_I5L; 1.
DR   ProDom; PD012198; Pox_I5; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   78 AA;  8779 MW;  AB8D48FF366D4D50 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 78;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
```

```
Db         42 RSLTF 46


RESULT 43
CATR1_HUMAN
ID   CATR1_HUMAN    STANDARD;      PRT;    79 AA.
AC   Q13166;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   CATR tumorigenic conversion 1 protein (CATR1.3).
GN   Name=CATR1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Carcinoma;
RX   MEDLINE=95327656; PubMed=7604004;
RA   Li D., Noyes I., Shuler C., Milo G.E.;
RT   "Cloning and sequencing of CATR1.3, a human gene associated with
RT   tumorigenic conversion.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:6409-6413(1995).
CC   -!- DEVELOPMENTAL STAGE: Associated with tumorigenic conversion.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U25433; -; NOT_ANNOTATED_CDS; mRNA.
DR   PIR; I38991; I38991.
DR   Ensembl; ENSG00000198950; Homo sapiens.
DR   HGNC; HGNC:1525; CATR1.
DR   MIM; 600676; -.
SQ   SEQUENCE   79 AA;  9224 MW;  BC3667C059114CF3 CRC64;

  Query Match             100.0%;  Score 15;  DB 1;  Length 79;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         42 RALTF 46


RESULT 44
Q9I2D4_PSEAE
ID   Q9I2D4_PSEAE PRELIMINARY;      PRT;    79 AA.
AC   Q9I2D4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PA1970;
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004623; AAG05358.1; -; Genomic_DNA.
DR   PIR; B83400; B83400.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   79 AA;  8661 MW;  E8900BF84ADCBE70 CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 79;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         44 RSLSF 48


RESULT 45
O14280_SCHPO
ID   O14280_SCHPO PRELIMINARY;      PRT;    80 AA.
AC   O14280;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   SPAC8C9.11 protein.
GN   ORFNames=SPAC8C9.11;
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,
RA   Sgouros J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K.D., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K.L., Murphy L.D., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K.M., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M.N., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J.R., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Duesterhoeft A., Fritzc C., Holzer E., Moestl D.,
RA   Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,
RA   Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
DR   EMBL; Z99168; CAB16299.1; -; Genomic_DNA.
DR   PIR; T39148; T39148.
DR   HSSP; Q8BGS2; 1V9J.
DR   GeneDB_Spombe; SPAC8C9.11; -.
DR   GO; GO:0030528; F:transcription regulator activity; IEA.
DR   InterPro; IPR002634; BolA.
DR   Pfam; PF01722; BolA; 1.
KW   Complete proteome.
SQ   SEQUENCE   80 AA;  9190 MW;  9031CFB15A876B7B CRC64;

  Query Match             100.0%;  Score 15;  DB 2;  Length 80;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          4 RXLXF 8
              | | |
Db         72 RTLSF 76


RESULT 46
Q6ZH28_ORYSA
ID   Q6ZH28_ORYSA PRELIMINARY;      PRT;    82 AA.
AC   Q6ZH28;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein OJ1399_H05.14.
GN   Name=OJ1399_H05.14;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP004090; BAD07633.1; -; Genomic_DNA.
DR   Gramene; Q6ZH28; -.
KW   Hypothetical protein.
SQ   SEQUENCE   82 AA;  8492 MW;  00BDA1A4341F7BF1 CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 82;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         33 RSLAF 37


RESULT 47
Q5LFH2_BACFN
ID   Q5LFH2_BACFN PRELIMINARY;      PRT;    82 AA.
AC   Q5LFH2;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=BF1411;
OS   Bacteroides fragilis (strain ATCC 25285 / NCTC 9343).
OC   Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC   Bacteroidaceae; Bacteroides.
OX   NCBI_TaxID=272559;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15746427; DOI=10.1126/science.1107008;
RA   Cerdeno-Tarraga A.-M., Patrick S., Crossman L.C., Blakely G.,
RA   Abratt V., Lennard N., Poxton I., Duerden B., Harris B., Quail M.A.,
RA   Barron A., Clark L., Corton C., Doggett J., Holden M.T.G., Larke N.,
RA   Line A., Lord A., Norbertczak H., Ormond D., Price C.,
RA   Rabbinowitsch E., Woodward J., Barrell B.G., Parkhill J.;
RT   "Extensive DNA inversions in the B. fragilis genome control variable
RT   gene expression.";
RL   Science 307:1463-1465(2005).
DR   EMBL; CR626927; CAH07122.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   82 AA;  9719 MW;  8919E43065D2A0C1 CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 82;
  Best Local Similarity   60.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         28 RSLSF 32


RESULT 48
Q72HW8_THET2
ID   Q72HW8_THET2 PRELIMINARY;      PRT;    83 AA.
AC   Q72HW8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=TTC1364;
OS   Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=262724;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15064768; DOI=10.1038/nbt956;
RA   Henne A., Brueggemann H., Raasch C., Wiezer A., Hartsch T.,
RA   Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA   Jacobi C., Starkuviene V., Schlenczeck S., Dencker S., Huber R.,
RA   Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT   "The genome sequence of the extreme thermophile Thermus
RT   thermophilus.";
RL   Nat. Biotechnol. 22:547-553(2004).
DR   EMBL; AE017305; AAS81706.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   83 AA;  9672 MW;  5BBA95D0121BFFAF CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 83;
  Best Local Similarity   60.0%;  Pred. No. 2e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         28 RALAF 32


RESULT 49
Q6M9S7_PARUW
ID   Q6M9S7_PARUW PRELIMINARY;      PRT;    83 AA.
AC   Q6M9S7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=pc1948;
OS   Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).
OC   Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.
OX   NCBI_TaxID=264201;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15073324;
RA   Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA   Fartmann B., Brandt P., Nyakatura G.J., Droege M., Frishman D.,
RA   Rattei T., Mewes H.-W., Wagner M.;
RT   "Illuminating the evolutionary history of chlamydiae.";
RL   Science 304:728-730(2004).
DR   EMBL; BX908798; CAF24672.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   83 AA;  9855 MW;  35F7CC84585D7B99 CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 83;
  Best Local Similarity   60.0%;  Pred. No. 2e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         57 RSLAF 61


RESULT 50
Q4RRI6_TETNG
```

```
ID   Q4RRI6_TETNG PRELIMINARY;      PRT;    84 AA.
AC   Q4RRI6;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 16 SCAF15002, whole genome shotgun sequence.
GN   ORFNames=GSTENG00030155001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01015002; CAG08996.1; -; Genomic_DNA.
SQ   SEQUENCE   84 AA;  9588 MW;  47E2E9B30EA67D5F CRC64;

  Query Match            100.0%;  Score 15;  DB 2;  Length 84;
  Best Local Similarity   60.0%;  Pred. No. 2e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          6 RTLSF 10


Search completed: May  4, 2006, 06:09:43
Job time : 237 secs
```

Chism, B.
09/726479
Seq. ID 2

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: May 4, 2006, 06:02:18 ; Search time 185 Seconds
(without alignments)
19.000 Million cell updates/sec

Title: US-09-726-470A-2
Perfect score: 15
Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15	100.0	5	9	ADZ71692	Adz71692 p21-deriv
2	15	100.0	5	9	ADZ72055	Adz72055 p21-deriv
3	15	100.0	10	7	ADM37109	Adm37109 Mutant Bc
4	15	100.0	10	7	ADM53404	Adm53404 Bcl-2 mut
5	15	100.0	11	7	ADE50790	Ade50790 Protein k
6	15	100.0	11	9	ADY82927	Ady82927 Protein k
7	15	100.0	11	9	ADZ80906	Adz80906 Amino aci
8	15	100.0	12	9	ADY82979	Ady82979 Protein k
9	15	100.0	13	7	ADE50791	Ade50791 Protein k
10	15	100.0	15	2	AAR30969	Aar30969 Jacalin f
11	15	100.0	17	6	ABP83159	Abp83159 G protein
12	15	100.0	18	6	ABP83165	Abp83165 G protein
13	15	100.0	19	8	ADR31537	Adr31537 Bcl-2 mut
14	15	100.0	20	5	ABJ10879	Abj10879 Murine rh
15	15	100.0	20	6	ABP83164	Abp83164 G protein
16	15	100.0	20	9	ADW52410	Human PL
17	15	100.0	21	2	AAR51378	Aar51378 J alpha s
18	15	100.0	21	8	ABO60589	Abo60589 Human gen
19	15	100.0	27	3	AAY78753	Aay78753 Hypervari
20	15	100.0	27	5	AAE18094	Aae18094 Biotinyla
21	15	100.0	28	3	AAB38998	Aab38998 Human sec
22	15	100.0	30	2	AAY02875	Aay02875 Fragment
23	15	100.0	30	7	ADA07705	Ada07705 Human sec
24	15	100.0	30	8	ADN41296	Adn41296 Novel hum
25	15	100.0	31	2	AAW45039	Aaw45039 Immunomod
26	15	100.0	31	2	AAY09486	Aay09486 Immunoact
27	15	100.0	32	9	ADV99762	Adv99762 Glucanase
28	15	100.0	33	3	AAB44770	Aab44770 Human sec
29	15	100.0	33	4	AAM32187	Aam32187 Peptide #
30	15	100.0	33	4	AAM60971	Aam60971 Human bra
31	15	100.0	33	4	ABG55401	Abg55401 Human liv
32	15	100.0	33	4	ABG53589	Abg53589 Human liv
33	15	100.0	33	5	ABG43538	Abg43538 Human pep
34	15	100.0	33	5	ABG41720	Abg41720 Human pep
35	15	100.0	35	3	AAB44347	Aab44347 Human sec
36	15	100.0	35	8	ABO56517	Abo56517 Human gen
37	15	100.0	36	3	AAB63066	Aab63066 Human sec
38	15	100.0	38	5	ABG68830	Abg68830 Cytochrom
39	15	100.0	40	3	AAG15201	Aag15201 Arabidops
40	15	100.0	40	7	ADD90505	Add90505 Novel hum
41	15	100.0	40	7	ADG90324	Adg90324 Human sec
42	15	100.0	40	8	ABO57531	Abo57531 Human gen
43	15	100.0	40	9	ADY25664	Ady25664 Novel hum
44	15	100.0	43	2	AAY25813	Aay25813 Human sec
45	15	100.0	45	4	AAU86594	Aau86594 Novel hum
46	15	100.0	45	7	ADB59928	Adb59928 Connectiv
47	15	100.0	45	8	ADN42941	Adn42941 Human sec
48	15	100.0	49	4	ABG26048	Abg26048 Novel hum
49	15	100.0	49	8	ADS98736	Ads98736 Protein f
50	15	100.0	50	4	AAM86216	Aam86216 Human imm
51	15	100.0	50	6	ABU07196	Abu07196 Maize SSI
52	15	100.0	50	6	ABU07203	Abu07203 Maize SSI
53	15	100.0	50	8	ADS07372	Ads07372 Staphyloc
54	15	100.0	53	5	ABG66860	Abg66860 Human pro
55	15	100.0	56	4	AAU67825	Aau67825 Propionib
56	15	100.0	56	5	ABP07362	Abp07362 Human ORF
57	15	100.0	56	5	ABG63221	Abg63221 Human pro
58	15	100.0	56	5	ABP64049	Abp64049 Human ORF
59	15	100.0	56	6	ABM64344	Abm64344 Propionib
60	15	100.0	57	5	ABP34657	Abp34657 Human ORF
61	15	100.0	58	3	AAY58437	Aay58437 Staphyloc
62	15	100.0	58	4	AAM38400	Aam38400 Peptide #
63	15	100.0	58	4	ABG26045	Abg26045 Novel hum
64	15	100.0	58	4	AAB69510	Aab69510 Staphyloc
65	15	100.0	58	5	ABG47169	Abg47169 Human pep
66	15	100.0	59	1	AAP91536	Aap91536 Modified
67	15	100.0	60	4	AAG99889	Aag99889 ERA bindi
68	15	100.0	60	7	ADC00737	Adc00737 Enterohae
69	15	100.0	60	8	ADX92123	Adx92123 Plant ful
70	15	100.0	61	2	AAY59684	Aay59684 Secreted
71	15	100.0	61	7	ADM77834	Adm77834 Human pro
72	15	100.0	61	8	ADP19515	Adp19515 Human sec
73	15	100.0	61	9	ADZ89294	Adz89294 Secreted
74	15	100.0	62	5	ABP33448	Abp33448 Human ORF
75	15	100.0	62	7	ADC94540	Adc94540 E. faeciu
76	15	100.0	63	4	AAM92859	Aam92859 Human dig
77	15	100.0	63	4	AAU22675	Aau22675 Novel hum
78	15	100.0	63	5	ABP06349	Abp06349 Human ORF
79	15	100.0	63	5	ABB49639	Abb49639 Listeria
80	15	100.0	63	5	ABJ10354	Abj10354 Human bre
81	15	100.0	63	7	ADB32515	Adb32515 Human nov
82	15	100.0	65	4	AAM17297	Aam17297 Peptide #
83	15	100.0	65	4	AAM04979	Aam04979 Peptide #
84	15	100.0	65	4	ABB17357	Abb17357 Human ner
85	15	100.0	65	5	ABP05290	Abp05290 Human ORF
86	15	100.0	66	4	ABB11901	Abb11901 Human cyt
87	15	100.0	66	8	ADS11950	Ads11950 Human the
88	15	100.0	66	8	ADS11954	Ads11954 Human the
89	15	100.0	66	8	ADS11953	Ads11953 Human the
90	15	100.0	66	8	ADS11951	Ads11951 Human the
91	15	100.0	66	8	ADS11952	Ads11952 Human the
92	15	100.0	67	3	AAY52246	Aay52246 M. pneumo
93	15	100.0	67	3	AAY52245	Aay52245 M. genita
94	15	100.0	67	3	AAY52239	Aay52239 E. coli y
95	15	100.0	67	4	AAM92271	Aam92271 Human dig
96	15	100.0	68	4	AAU62501	Aau62501 Propionib
97	15	100.0	68	5	ABP00761	Abp00761 Human ORF

98	15	100.0	68	6	ABM59020	Abm59020 Propionib
99	15	100.0	69	4	AAU17820	Aau17820 Novel hum
100	15	100.0	69	7	ADG41200	Adg41200 Human res

ALIGNMENTS

RESULT 1
ADZ71692
ID ADZ71692 standard; peptide; 5 AA.
XX
AC ADZ71692;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #277.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 20; Page 101; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 15; DB 9; Length 5;
Best Local Similarity 60.0%; Pred. No. 2e+06;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 RXLXF 8
              | | |
Db          1 RSLAF 5
```

RESULT 2
ADZ72055
ID ADZ72055 standard; peptide; 5 AA.
XX
AC ADZ72055;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #640.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Example 27; Page 83; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;

```
CC   each of B or D is independently an amino acid residue selected from
CC   arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC   isoleucine or alanine; C is a natural or unnatural amino acid residue
CC   having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC   containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC   or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC   Also described are: a pharmaceutical composition comprising the peptide
CC   admixed with a diluent, an excipient or a carrier; an assay for
CC   identifying candidate substances capable of binding to a cyclin
CC   associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC   assay for identifying compounds that interact a cyclin or a cyclin when
CC   complexed with the physiologically relevant CDK; and a method of using a
CC   cyclin in a drug screening assay. The assay for identifying candidate
CC   substances capable of binding to a cyclin associated with a G1 control
CC   CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC   a peptide as defined above, the cyclin, the CDK and the candidate
CC   substance, under conditions where, in the absence of the candidate
CC   substance being an inhibitor of interaction of the cyclin/CDK
CC   interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC   any change in the expected binding of the peptide and the cyclin. The
CC   assay for identifying compounds that interact a cyclin or a cyclin when
CC   complexed with the physiologically relevant CDK comprises: incubating a
CC   candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC   and detecting binding of either the candidate compound or the peptide
CC   with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC   comprises use of a three-dimensional model of a cyclin and a candidate
CC   compound. At least one of the assay components is bound to a solid phase.
CC   The peptidomimetic is labeled such as to emit a signal when bound to the
CC   cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC   peptide. One of the assay components is labeled with a fluorescence
CC   emitter and the signal is detected using fluorescence polarization
CC   techniques. Using a cyclin in a drug screening assay comprises: selecting
CC   a candidate compound by performing rational drug design with a three-
CC   dimensional model of the cyclin, where the selecting is performed in
CC   conjunction with computer modeling; contacting the candidate compound
CC   with the cyclin; and detecting the binding of the candidate compound for
CC   the cyclin groove. A potential drug is selected on the basis of its
CC   having a greater affinity for the cyclin groove than that of the peptide.
CC   The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC   cell cycle-related apoptosis, suppression of E2F transcription factor,
CC   hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC   effects. The peptide is useful in preparing a medicament for treating a
CC   proliferative disorder, e.g., cancer. The present sequence represents a
CC   p21-derived peptide of the invention.
XX
SQ   Sequence 5 AA;

  Query Match             100.0%;  Score 15;  DB 9;  Length 5;
  Best Local Similarity   60.0%;  Pred. No. 2e+06;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          1 RSLAF 5


RESULT 3
ADM37109
ID   ADM37109 standard; peptide; 10 AA.
XX
AC   ADM37109;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Mutant Bcl-2 81-90 #1.
XX
KW   Bcl-2; AKT kinase; kinase; phosphorylation substrate;
KW   kinase recognition domain; E3 binding region; mutant; mutein.
XX
OS   Synthetic.
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "Wild-type Ala substituted by Leu"
FT   Misc-difference 4
FT                   /note= "Wild-type Ala substituted by Arg"
FT   Misc-difference 8
FT                   /note= "Wild-type Pro substituted by Phe"
XX
PN   US2003170611-A1.
XX
PD   11-SEP-2003.
XX
PF   09-MAR-2002; 2002US-00093840.
XX
PR   09-MAR-2002; 2002US-00093840.
XX
PA   (CARD/) CARDONE M H.
PA   (YAFF/) YAFFE M.
XX
PI   Cardone MH,  Yaffe M;
XX
DR   WPI; 2003-863752/80.
XX
PT   Identifying a molecule capable of modulating a kinase activity in situ
PT   using a substrate with an altered kinase recognition domain and an
PT   associated label is useful in drug discovery.
XX
PS   Disclosure; Page 8; 25pp; English.
XX
CC   The invention relates to identifying a molecule capable of modulating a
CC   kinase activity in situ, comprising exposing a candidate molecule to a
CC   cell comprising a phosphorylation substrate associated with a detectable
CC   label and having a kinase recognition domain altered to be recognised by
CC   a kinase that does not recognise the substrate in its unaltered state,
CC   and determining whether the candidate molecule causes a change in an
CC   expression of the label. Also included are a molecule capable of
CC   modulating activity of at least one kinase in situ identified by the
CC   claimed method, a fusion protein (comprising an E3 binding region, a
CC   kinase recognition domain and a green fluorescent protein, where the
CC   kinase recognition domain is the domain of beta-catenin, HIV protein VPU,
CC   p27, Bcl-2 or c-Jun), a fusion protein (comprising an E3 binding region,
CC   a kinase recognition domain and an enzyme capable of producing a
CC   detectable enzymatic product, where the kinase recognition domain is the
CC   domain of beta-catenin, HIV protein VPU, p27, Bcl-2 or c-Jun), an
CC   isolated genetic molecule encoding one of the above fusion proteins, a
CC   vector capable of expressing the above genetic molecule and a cell
CC   transfected with the above vector. Expression of the label requires
CC   phosphorylation of the phosphorylation substrate by the kinase. The
CC   kinase recognition domain is altered to include a consensus recognition
CC   motif for the kinase (e.g. AKT kinase). The invention is useful to study
CC   kinase activity in situ and to screen for molecules that modulate kinase
CC   activities in situ, for example in drug discovery. The invention allows
CC   for information on multiple kinases to be provided simultaneously, which
CC   prior art does not provide. The present sequence is a mutated region
CC   (ERK1/2 binding region) of Bcl-2, which is mutated to an AKT or
CC   CHGK1/2(not defined) phosphorylation site.
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 15;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 3.1e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          4 RALSF 8


RESULT 4
ADM53404
ID   ADM53404 standard; peptide; 10 AA.
XX
AC   ADM53404;
XX
```

DT 03-JUN-2004 (first entry)
XX
DE Bcl-2 mutant CHK2 substrate domain.
XX
KW kinase activity modulator; signaling enzyme; drug design; anti-tumour;
KW anti-inflammation; anti-ischaemia; bcl-2; CHK2; mutant; mutein.
XX
OS Unidentified.
OS Synthetic.
XX
PN US2003170737-A1.
XX
PD 11-SEP-2003.
XX
PF 09-MAR-2002; 2002US-00093945.
XX
PR 09-MAR-2002; 2002US-00093945.
XX
PA (CARD/) CARDONE M H.
PA (YAFF/) YAFFE M.
XX
PI Cardone MH, Yaffe M;
XX
DR WPI; 2003-787546/74.
XX
PT A method for identifying a molecule capable of modulating a kinase
PT activity in situ and is useful in drug design to screen for molecules
PT that are candidates for anti-tumor, anti-inflammation, and anti-ischemia
PT therapy.
XX
PS Disclosure; Page 8; 25pp; English.
XX
CC The invention describes a method of identifying a molecule capable of
CC modulating a kinase activity in situ, comprising exposing a candidate
CC molecule to a cell comprising a signaling enzyme altered so as to bind a
CC phosphorylation substrate, where the substrate is associated with
CC detectable label and the binding is regulated by a kinase, and
CC determining whether the candidate causes a change in expression of the
CC label. Also described are: a molecule capable of modulating a kinase
CC activity in situ identified by the claimed method; a fusion protein
CC comprising a genetically altered signaling enzyme and a label, where the
CC alteration produces an adapter molecule in the signaling enzyme capable
CC of binding to a phosphorylation substrate that the enzyme does not bind
CC in its unaltered state, where the binding is regulated by a kinase; an
CC isolated genetic molecule encoding the above fusion protein; a vector
CC capable of expressing the above genetic molecule; and a cell transfected
CC with the above vector. The invention is useful in drug design to screen
CC for molecules that are candidates for anti-tumour, anti-inflammation, and
CC anti-ischaemia therapy. This is the amino acid sequence of a Bcl-2 ERK1
CC substrate mutant that has been altered to change the kinase consensus
CC sequence from ERK1 to CHK2.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 15; DB 7; Length 10;
Best Local Similarity 60.0%; Pred. No. 3.1e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 4 RALSF 8

RESULT 5
ADE50790
ID ADE50790 standard; peptide; 11 AA.
XX
AC ADE50790;
XX
DT 29-JAN-2004 (first entry)
XX
DE Protein kinase substrate polypeptide-related peptide 12.
XX
KW protein kinase substrate polypeptide; protein kinase activity;
KW threonine protein kinase; serine protein kinase.
XX
OS Synthetic.
OS Unidentified.
XX
PN WO2003087400-A1.
XX
PD 23-OCT-2003.
XX
PF 26-MAR-2003; 2003WO-GB001286.
XX
PR 09-APR-2002; 2002GB-00008104.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Armstrong C, Cohen P;
XX
DR WPI; 2003-845342/78.
XX
PT New kit for assaying protein kinase activity and in screening for
PT modulators of protein kinase activity, comprises protein kinase substrate
PT polypeptides comprising a specificity conferring portion and a
PT phosphorylatable portion.
XX
PS Disclosure; Fig 2; 48pp; English.
XX
CC This invention relates to a novel kit of parts comprising two or more
CC protein kinase substrate polypeptides. The kit and methods are useful in
CC assaying protein kinase activity and in screening for modulators of
CC protein kinase activity, in particular serine/threonine protein kinases.
CC The polypeptide or phosphorylated polypeptide or the antibody is also
CC used in an assay of protein kinase activity. The present sequence is that
CC of a peptide which is related to the invention.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 15; DB 7; Length 11;
Best Local Similarity 60.0%; Pred. No. 3.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 3 RTLSF 7

RESULT 6
ADY82927
ID ADY82927 standard; peptide; 11 AA.
XX
AC ADY82927;
XX
DT 02-JUN-2005 (first entry)
XX
DE Protein kinase C (PKC) substrate peptide - SEQ ID 200.
XX
KW enzyme engineering; phosphorylation; protein kinase C; PKC; substrate.
XX
OS Synthetic.
XX
PN US2005064507-A1.
XX
PD 24-MAR-2005.
XX
PF 11-SEP-2003; 2003US-00660370.
XX
PR 11-SEP-2003; 2003US-00660370.
XX
PA (SHAW/) SHAW J S.
XX
PI Shaw JS;
XX
DR WPI; 2005-252669/26.

```
XX
PT   Test set useful for characterizing substrate specificities of kinases,
PT   comprises two peptide pools, in which every peptide in each of the
PT   peptide pools has phosphorylatable, query, anchor and degenerate amino
PT   acid positions.
XX
PS   Claim 31; SEQ ID NO 200; 201pp; English.
XX
CC   The invention comprises a test set for characterizing substrate
CC   specificities of kinases, consisting of two peptide pools, in which every
CC   peptide in each of the peptide pools contains one phosphorylatable amino
CC   acid position, one query amino acid position, at least one anchor amino
CC   acid position, and at least one degenerate amino acid position. The test
CC   set of the invention is useful to determine the spectrum of peptidyl
CC   sequences that are phosphorylated by a kinase. The present amino acid
CC   sequence represents a peptide that was used as a substrate for a protein
CC   kinase C (PKC).
XX
SQ   Sequence 11 AA;

  Query Match           100.0%;  Score 15;  DB 9;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 3.4e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          5 RALSF 9


RESULT 7
ADZ80906
ID   ADZ80906 standard; peptide; 11 AA.
XX
AC   ADZ80906;
XX
DT   14-JUL-2005  (first entry)
XX
DE   Amino acid sequence of OAS1 gene conserved region #4.
XX
KW   oligoadenylate synthetase gene; OAS1; viral infection;
KW   hepatitis C virus infection; antiinflammatory; hepatotropic;
KW   gastrointestinal disease; infection; cytostatic; virucide; antidiabetic;
KW   endocrine disease; gastrointestinal disease; metabolic disorder;
KW   neuroleptic; psychiatric disorder; neoplasm; andrology;
KW   genitourinary disease; gene therapy; diabetes mellitus; schizophrenia;
KW   cancer; prostate tumor; interferon-based treatment.
XX
OS   Homo sapiens.
XX
PN   WO2005040428-A2.
XX
PD   06-MAY-2005.
XX
PF   22-OCT-2004; 2004WO-US035284.
XX
PR   23-OCT-2003; 2003US-0513888P.
PR   06-FEB-2004; 2004US-0542373P.
PR   19-MAR-2004; 2004US-0554758P.
PR   08-APR-2004; 2004US-0560524P.
PR   09-JUN-2004; 2004US-0578323P.
PR   28-JUN-2004; 2004US-0583503P.
PR   26-AUG-2004; 2004US-0605243P.
XX
PA   (ILLU-) ILLUMIGEN BIOSCIENCES INC.
XX
PI   Iadonato SP,  Magness CL,  Rosenberg G,  Scherer CA;
XX
DR   WPI; 2005-333517/34.
XX
PT   Identifying an oligoadenylate synthetase 1 gene mutation, useful for
PT   treating, e.g. viral infection, cancer or diabetes, comprises detecting
PT   in a nucleic acid sample, the presence of an oligoadenylate synthetase 1
PT   point mutation.
XX
PS   Claim 3; SEQ ID NO 78; 111pp; English.
XX
CC   The specification describes a method for identifying an oligoadenylate
CC   synthetase gene (OAS1) mutation. The method comprises detecting in a
CC   nucleic acid sample the presence of an OAS1 point mutation, preferably
CC   substitution of a non-reference nucleotide for a reference nucleotide at
CC   nucleotide position 2135728, 2135749, 2135978, 2144072, 2144088, 2144116,
CC   2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, and
CC   2156638 of reference sequence ADZ80847.The mutations correlate with
CC   resistance of humans to viral infection, such as Hepatitis C. The method
CC   of the invention is useful for identifying an OAS1 mutation, and for
CC   identifying susceptibility to viral infection, hepatitis C infection,
CC   predisposition to diabetes mellitus or schizophrenia in humans,
CC   susceptibility to cancer such as prostate cancer, or for identifying a
CC   patient responsiveness to therapeutic treatments for viral infection,
CC   where the therapeutic treatment is interferon-based and the patient
CC   response is measured by sustained viral clearance. Peptides ADZ80903-
CC   ADZ80912 represent conserved regions of the OAS1 gene.
XX
SQ   Sequence 11 AA;

  Query Match           100.0%;  Score 15;  DB 9;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 3.4e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          3 RALSF 7


RESULT 8
ADY82979
ID   ADY82979 standard; peptide; 12 AA.
XX
AC   ADY82979;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Protein kinase C (PKC) substrate peptide - SEQ ID 321.
XX
KW   enzyme engineering; phosphorylation; protein kinase C; PKC; substrate.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 8
FT                   /note= "Phosphorylated Thr"
XX
PN   US2005064507-A1.
XX
PD   24-MAR-2005.
XX
PF   11-SEP-2003; 2003US-00660370.
XX
PR   11-SEP-2003; 2003US-00660370.
XX
PA   (SHAW/) SHAW J S.
XX
PI   Shaw JS;
XX
DR   WPI; 2005-252669/26.
XX
PT   Test set useful for characterizing substrate specificities of kinases,
PT   comprises two peptide pools, in which every peptide in each of the
PT   peptide pools has phosphorylatable, query, anchor and degenerate amino
PT   acid positions.
XX
PS   Claim 41; SEQ ID NO 321; 201pp; English.
XX
CC   The invention comprises a test set for characterizing substrate
CC   specificities of kinases, consisting of two peptide pools, in which every
CC   peptide in each of the peptide pools contains one phosphorylatable amino
```

```
CC   acid position, one query amino acid position, at least one anchor amino
CC   acid position, and at least one degenerate amino acid position. The test
CC   set of the invention is useful to determine the spectrum of peptidyl
CC   sequences that are phosphorylated by a kinase. The present amino acid
CC   sequence represents a peptide that was used as a substrate for a protein
CC   kinase C (PKC)..
XX
SQ   Sequence 12 AA;

  Query Match             100.0%;  Score 15;  DB 9;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.7e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          1 RALSF 5


RESULT 9
ADE50791
ID   ADE50791 standard; peptide; 13 AA.
XX
AC   ADE50791;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Protein kinase substrate polypeptide-related peptide 13.
XX
KW   protein kinase substrate polypeptide; protein kinase activity;
KW   threonine protein kinase; serine protein kinase.
XX
OS   Synthetic.
OS   Unidentified.
XX
PN   WO2003087400-A1.
XX
PD   23-OCT-2003.
XX
PF   26-MAR-2003; 2003WO-GB001286.
XX
PR   09-APR-2002; 2002GB-00008104.
XX
PA   (UYDU-) UNIV DUNDEE.
XX
PI   Armstrong C,  Cohen P;
XX
DR   WPI; 2003-845342/78.
XX
PT   New kit for assaying protein kinase activity and in screening for
PT   modulators of protein kinase activity, comprises protein kinase substrate
PT   polypeptides comprising a specificity conferring portion and a
PT   phosphorylatable portion.
XX
PS   Disclosure; Fig 2; 48pp; English.
XX
CC   This invention relates to a novel kit of parts comprising two or more
CC   protein kinase substrate polypeptides. The kit and methods are useful in
CC   assaying protein kinase activity and in screening for modulators of
CC   protein kinase activity, in particular serine/threonine protein kinases.
CC   The polypeptide or phosphorylated polypeptide or the antibody is also
CC   used in an assay of protein kinase activity. The present sequence is that
CC   of a peptide which is related to the invention.
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 15;  DB 7;  Length 13;
  Best Local Similarity   60.0%;  Pred. No. 4e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          5 RTLSF 9


RESULT 10
AAR30969
ID   AAR30969 standard; peptide; 15 AA.
XX
AC   AAR30969;
XX
DT   25-MAR-2003  (revised)
DT   07-MAY-1993  (first entry)
XX
DE   Jacalin fragment which interacts with CD4 receptor.
XX
KW   Jackfruit; human immunodeficiency virus; HIV-1; gp120; mitogenic lectin.
XX
OS   Artocarpus heterophyllus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "uncharged hydrophilic amino acid - opt. absent"
XX
PN   WO9222574-A1.
XX
PD   23-DEC-1992.
XX
PF   05-JUN-1992;   92WO-FR000510.
XX
PR   10-JUN-1991;   91FR-00007041.
PR   31-JAN-1992;   92FR-00001127.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MED.
XX
PI   Favero J,  Dornand J,  Corbeau P,  Devaux C,  Nicolas M,  Liautard J;
XX
DR   WPI; 1993-018076/02.
XX
PT   Jacalin and its new peptide fragments for treating HIV - interacting with
PT   the CD4 receptor and specifically preventing infection of lymphocytes.
XX
PS   Claim 7; Page 14; 27pp; French.
XX
CC   This fragment of Jacalin interacts with the CD4 receptor and is
CC   homologous with the sequence of the HIV protein gp120. The peptide and
CC   other peptides with biological activity equivalent to that of Jacalin are
CC   useful in treatment of diseases caused by HIV. They specifically inhibit
CC   infection of lymphocytes by HIV, do not affect normal lymphocyte function
CC   and (unlike Jacalin itself) do not agglutinate cells. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 15;  DB 2;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          5 RSLTF 9


RESULT 11
ABP83159
ID   ABP83159 standard; peptide; 17 AA.
XX
AC   ABP83159;
XX
DT   04-MAR-2003  (first entry)
XX
DE   G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1832.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
```

KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
PS Claim 1; Fig 2; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 15; DB 6; Length 17;
Best Local Similarity 60.0%; Pred. No. 5.1e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 9 RSLSF 13

RESULT 12
ABP83165
ID ABP83165 standard; peptide; 18 AA.
XX
AC ABP83165;
XX
DT 04-MAR-2003 (first entry)
XX
DE G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1838.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
PS Claim 1; Fig 2; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 15; DB 6; Length 18;
Best Local Similarity 60.0%; Pred. No. 5.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8

```
                | | |
Db          1 RSLSF 5


RESULT 13
ADR31537
ID   ADR31537 standard; peptide; 19 AA.
XX
AC   ADR31537;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Bcl-2 mutant substrate peptide #1.
XX
KW   Drug designing; clinical application; anti-inflammatory; anti-tumour;
KW   immune response; anti-ischaemic; artherosclerosis; therapy; Bcl-2;
KW   mutant; mutein.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "Wild type Thr substituted with Leu"
FT   Misc-difference 4
FT                   /note= "Wild type Pro substituted with Arg"
FT   Misc-difference 8
FT                   /note= "Wild type Pro substituted with Phe"
FT   Misc-difference 11
FT                   /note= "Wild type Ala substituted with Leu"
FT   Misc-difference 13
FT                   /note= "Wild type Pro substituted with Arg"
FT   Misc-difference 17
FT                   /note= "Wild type Pro substituted with Phe"
XX
PN   US2004157272-A1.
XX
PD   12-AUG-2004.
XX
PF   03-FEB-2004; 2004US-00771035.
XX
PR   09-MAR-2002; 2002US-00093945.
XX
PA   (MERR-) MERRIMACK PHARM INC.
XX
PI   Cardone MH,  Yaffe M;
XX
DR   WPI; 2004-580267/56.
XX
PT   Identifying molecule capable of modulating kinase activity in situ, by
PT   exposing candidate molecule to cell comprising signaling enzyme altered
PT   to bind substrate having label and determining whether molecule changes
PT   expression of label.
XX
PS   Disclosure; Page 8; 25pp; English.
XX
CC   The invention relates to a method for identifying molecule capable of
CC   modulating kinase activity in situ. The method involves exposing
CC   candidate molecule to cell having signalling enzyme altered to bind
CC   phosphorylation substrate associated with detectable label where the
CC   binding between altered signalling enzyme and substrate is regulated by a
CC   kinase and determining if the candidate molecule cause a change in
CC   expression of label. The method is useful for testing and designing drugs
CC   with various clinical application e.g. anti-inflammatory candidate
CC   molecules. The invention is useful in the field of anti-tumour
CC   therapeutics and immune response regulating drugs. The method is also
CC   useful for developing anti-ischaemic drugs that are useful for treating
CC   artherosclerosis. The present sequence is Bcl-2 mutant substrate peptide.
CC   This sequence is used to illustrate the method of the invention.
XX
SQ   Sequence 19 AA;

  Query Match             100.0%;  Score 15;  DB 8;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 5.7e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         13 RALSF 17


RESULT 14
ABJ10879
ID   ABJ10879 standard; peptide; 20 AA.
XX
AC   ABJ10879;
XX
DT   02-DEC-2002  (first entry)
XX
DE   Murine rhophilin protein SEQ ID NO: 164.
XX
KW   Human; PAL-18; cancer; chromosome 1q41; prostate cancer; colon cancer;
KW   breast cancer; cytostatic.
XX
OS   Mus musculus.
XX
PN   US2002106765-A1.
XX
PD   08-AUG-2002.
XX
PF   12-MAR-2001; 2001US-00804682.
XX
PR   10-MAR-2000; 2000US-0188586P.
XX
PA   (KIND/) KINDERS R J.
PA   (CORE/) COREY M J.
XX
PI   Kinders RJ,  Corey MJ;
XX
DR   WPI; 2002-697869/75.
XX
PT   New isolated PAL-18 polypeptide, useful for diagnosing, characterizing,
PT   and treating disease and in determining disease susceptibility.
XX
PS   Disclosure; Page 105; 150pp; English.
XX
CC   The present invention relates to human PAL-18 polypeptides and
CC   polynucleotides. The PAL-18 gene is found on chromosome 1q41. The
CC   sequences can be used to diagnose, monitor and treat cancers,
CC   particularly breast, colon and prostate cancers. The present sequence is
CC   a protein shown in the invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 15;  DB 5;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 6e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         15 RALAF 19


RESULT 15
ABP83164
ID   ABP83164 standard; peptide; 20 AA.
XX
AC   ABP83164;
XX
DT   04-MAR-2003  (first entry)
XX
DE   G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1837.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
```

```
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer.
XX
OS   Homo sapiens.
XX
PN   WO200261087-A2.
XX
PD   08-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US050107.
XX
PR   19-DEC-2000; 2000US-0257144P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Burmer GC,  Roush CL,  Brown JP;
XX
DR   WPI; 2003-046718/04.
XX
PT   New isolated antigenic peptides e.g., for G protein-coupled receptors
PT   (GPCR), useful for diagnosing and designing drugs for treating conditions
PT   in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT   autoimmune diseases.
XX
PS   Claim 1; Fig 2; 523pp; English.
XX
CC   The present invention describes antigenic peptides (I) comprising: (a)
CC   any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular G
CC   protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity or
CC   avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC   gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC   antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 15;  DB 6;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 6e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         14 RSLSF 18


RESULT 16
ADW52410
ID   ADW52410 standard; peptide; 20 AA.
XX
AC   ADW52410;
XX
DT   07-APR-2005  (first entry)
XX
DE   Human PL protein PDZ binding peptide #425.
XX
KW   antidepressant; cytostatic; uropathic; nootropic; vulnerary; vasotropic;
KW   neuroprotective; antidote; antiparkinsonian; ophthalmological;
KW   alpha-adrenergic receptor modulator; depression; apoptosis; migraine;
KW   Alzheimers disease; Parkinsons disease; glaucoma; drug dependence.
XX
OS   Homo sapiens.
XX
PN   US2005019841-A1.
XX
PD   27-JAN-2005.
XX
PF   14-OCT-2003; 2003US-00684796.
XX
PR   14-MAY-1999;   99US-0134114P.
PR   12-MAY-2000; 2000US-00570118.
PR   28-NOV-2000; 2000US-00724553.
PR   16-FEB-2001; 2001US-0269523P.
PR   03-AUG-2001; 2001US-0309841P.
PR   19-FEB-2002; 2002US-00080273.
PR   25-FEB-2002; 2002US-0360061P.
PR   02-AUG-2002; 2002WO-US024655.
PR   11-OCT-2002; 2002US-0418042P.
PR   14-NOV-2002; 2002US-0426212P.
XX
PA   (ARBO-) ARBOR VITA CORP.
XX
PI   Garman JD,  Lu PS;
XX
DR   WPI; 2005-131132/14.
XX
PT   New peptide that modulates binding between specific PDZ polypeptide and
PT   alpha-adrenergic receptor PL polypeptide, useful for treating depression,
PT   hypertrophy, Alzheimer's and Parkinson's.
XX
PS   Example 3; Page 96; 111pp; English.
XX
CC   The invention relates to a modulator (I) of binding between a specific
CC   PDZ polypeptide and an alpha-adrenergic receptor PL polypeptide, which is
CC   a peptide (P1) comprising at least 3 residues of a C-terminal sequence
CC   demonstrated to bind the target PDZ polypeptide, or a peptide mimetic
CC   (P2) of P1, or a small molecule having similar functional activity as the
CC   peptide with respect to PDZ polypeptide and PL polypeptide binding pair.
CC   Modulator of PDZ-PL interactions (claimed). (I) and PC1 are useful for
CC   treating a disorder such as depression, lower urinary tract storage,
CC   prostate apoptosis, hypertrophy, proliferation and migration of vascular
CC   smooth muscle after carotid injury, migraine, coronary flow reserve
CC   following stenting, Alzheimer's, Parkinson's, neuroprotection, glaucoma
CC   and opioid withdrawal. This sequence corresponds to peptide from a PL
CC   protein which binds a PDZ containing protein used in the invention.
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 15;  DB 9;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 6e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          1 RSLSF 5


RESULT 17
AAR51378
ID   AAR51378 standard; protein; 21 AA.
XX
AC   AAR51378;
XX
DT   25-MAR-2003  (revised)
```

DT 20-OCT-1994 (first entry)
XX
DE J alpha sequence (Va12.1/JaA?? usage).
XX
KW Rheumatoid arthritis; antibody; TCR; T cell receptor; lymphocyte;
KW expansion; complementary determining region; CDR3; antigen; MHC.
XX
OS Homo sapiens.
XX
PN WO9406823-A1.
XX
PD 31-MAR-1994.
XX
PF 14-SEP-1993; 93WO-US008644.
XX
PR 14-SEP-1992; 92US-00943418.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
XX
PI Brenner MB, Dersimonian H;
XX
DR WPI; 1994-118395/14.
XX
PT Treatment and prevention of rheumatoid arthritis - using peptides derived
PT from DQw2 or antibodies to DQw2 which block activation of T lymphocytes.
XX
PS Disclosure; Fig 3A; 57pp; English.
XX
CC To gain insight into the basis fo the Valpha12.1+T cell expansion in
CC rheumatoid arthritis, Valpha12.1 transcripts from positively selected
CC CD8+T cells were cloned and sequenced. In each of the three patients
CC analysed, distinct, repeated Valpha12.1 contg. sequences corresp. to
CC functional TCR alpha-chain transcripts were identified. All of the
CC repeated Valpha12.1+ T cell rearrangements in the 3 patients analysed use
CC either JalphaA1, JalphaA12 or JalphaA6, each of which encodes a unique
CC sequence at the 3' end of the Jalpha gene segment. This short stretch of
CC shared residues (pro-tyr) is predicted to contribute (or is immediately
CC adjacent) to the third complementary determining region (CDR3) and thus
CC may play a role in antigen or MHC recognition. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 15; DB 2; Length 21;
Best Local Similarity 60.0%; Pred. No. 6.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 6 RALTF 10

RESULT 18
ABO60589
ID ABO60589 standard; protein; 21 AA.
XX
AC ABO60589;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #6823.
XX
KW Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 45; SEQ ID NO 34223; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 15; DB 8; Length 21;
Best Local Similarity 60.0%; Pred. No. 6.3e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 9 RSLSF 13

RESULT 19
AAY78753
ID AAY78753 standard; peptide; 27 AA.
XX
AC AAY78753;
XX
DT 08-MAY-2000 (first entry)
XX
DE Hypervariable region 1 representative peptide sequence 9.

XX
KW Hepatitis C virus; envelope protein E2; hypervariable region 1; mimitope;
KW peptide library; treatment; prevent infection; antibody production.
XX
OS Hepatitis C virus.
XX
PN WO9960132-A1.
XX
PD 25-NOV-1999.
XX
PF 14-MAY-1999; 99WO-EP003344.
XX
PR 19-MAY-1998; 98GB-00010756.
XX
PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI Nicosia A, Lahm A, Tramontano A, Cortese R;
XX
DR WPI; 2000-126382/11.
XX
PT A new peptide library from hepatitis C virus, useful for production of
PT treatment for hepatitis C.
XX
PS Example; Page 73; 126pp; English.
XX
CC This sequence represents a peptide from the library of the invention. The
CC invention relates to a library of peptides which have an immunologically
CC reactive epitope of the hypervariable region 1 (HVR1) of envelope protein
CC 2 (E2) of hepatitis C virus. The peptides contained in the library
CC correspond to formulae given in the specification (see AAY78596-Y78598).
CC This sequence is included in a selection of a representative set of
CC natural HVR1 sequences. The peptides can be used in a method to select
CC antibodies which react with the HVR1 of E2 of hepatitis C virus, through
CC the selection of those antibodies which bind to the peptides. The
CC peptides from hepatitis C virus hypervariable region 1 of the envelope
CC protein E2 are used to produce a medicament for raising or increasing
CC levels of antibodies able to bind HCV (hepatitis C virus) HVR1 epitopes
CC in a mammal. The medicament is used to treat or prevent an HCV infection
XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 15; DB 3; Length 27;
Best Local Similarity 60.0%; Pred. No. 8.1e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 1 RTLSF 5

RESULT 20
AAE18094
ID AAE18094 standard; peptide; 27 AA.
XX
AC AAE18094;
XX
DT 07-MAY-2002 (first entry)
XX
DE Biotinylated hepatitis C virus region derived peptide (HVR), 272.
XX
KW Hepatitis C virus; HCV conjugate; immune response; therapeutic; virucide;
KW hepatotropic; antiinflammatory; HCV region derived peptide; HVR.
XX
OS Hepatitis C virus.
XX
PN WO200193804-A2.
XX
PD 13-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017302.
XX
PR 02-JUN-2000; 2000US-0209089P.
XX
PA (MERI) MERCK & CO INC.
XX
PI Conley AJ, Mckenna PM, Przysiecki CT, Keller PM;
XX
DR WPI; 2002-164292/21.
XX
PT Hepatitis C virus conjugate useful for inducing immune response in a
PT subject comprises a polypeptides or protein complex carrier and
PT immunogenic peptides covalently bonded to the carrier.
XX
PS Example 1; Page 22; 63pp; English.
XX
CC The patent discloses hepatitis C virus (HCV) conjugates able to induce an
CC immune response recognising different strains and variants of HCV. The
CC conjugates comprise a polypeptide or protein complex carrier and one or
CC more HCV mimotopes. Sequences of the invention are useful for inducing an
CC immune response in a subject e.g. human, chimpanzees, mice or horses.
CC They are also useful for the preparation of antisera, in
CC therapeutic/diagnostic applications to generate anti-HCV antibodies, for
CC detecting the presence of HCV in a subject and treating the subject
CC infected with HCV. The present sequence is biotinylated HCV region
CC derived peptide (HVR), 270
XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 15; DB 5; Length 27;
Best Local Similarity 60.0%; Pred. No. 8.1e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 1 RTLSF 5

RESULT 21
AAB38998
ID AAB38998 standard; peptide; 28 AA.
XX
AC AAB38998;
XX
DT 02-FEB-2001 (first entry)
XX
DE Human secreted peptide #20.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200056880-A1.
XX
PD 28-SEP-2000.
XX
PF 16-MAR-2000; 2000WO-US006781.
XX
PR 19-MAR-1999; 99US-0125363P.
PR 08-DEC-1999; 99US-0169617P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-602220/57.
DR N-PSDB; AAC59706.
XX
PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating disorders such as Parkinson's and
PT Alzheimer's diseases, cancers and infections.
XX
PS Claim 11; Page 380; 422pp; English.

```
XX
CC   Sequences AAB38971-B39020 represent the amino acid sequences of 50 human
CC   secreted proteins encoded by the genes AAC59679-C59728. The genes and
CC   proteins are useful for preventing, ameliorating or treating medical
CC   conditions, e.g. by protein or gene therapy. The genes are isolated from
CC   a range of human tissues disclosed in the specification. The nucleic
CC   acids, proteins, antibodies and (ant)agonists are useful in the
CC   diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC   immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections.
XX
SQ   Sequence 28 AA;

  Query Match             100.0%;  Score 15;  DB 3;  Length 28;
  Best Local Similarity   60.0%;  Pred. No. 8.4e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          9 RTLAF 13


RESULT 22
AAY02875
ID   AAY02875 standard; protein; 30 AA.
XX
AC   AAY02875;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Fragment of human secreted protein encoded by gene 85.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9902546-A1.
XX
PD   21-JAN-1999.
XX
PF   07-JUL-1998;   98WO-US013684.
XX
PR   08-JUL-1997;   97US-0051916P.
PR   08-JUL-1997;   97US-0051918P.
PR   08-JUL-1997;   97US-0051919P.
PR   08-JUL-1997;   97US-0051920P.
PR   08-JUL-1997;   97US-0051925P.
PR   08-JUL-1997;   97US-0051926P.
PR   08-JUL-1997;   97US-0051928P.
PR   08-JUL-1997;   97US-0051929P.
PR   08-JUL-1997;   97US-0051930P.
PR   08-JUL-1997;   97US-0051931P.
PR   08-JUL-1997;   97US-0051932P.
PR   08-JUL-1997;   97US-0052732P.
PR   08-JUL-1997;   97US-0052733P.
PR   08-JUL-1997;   97US-0052793P.
PR   08-JUL-1997;   97US-0052795P.
PR   08-JUL-1997;   97US-0052803P.
PR   18-AUG-1997;   97US-0055684P.
PR   18-AUG-1997;   97US-0055722P.
PR   18-AUG-1997;   97US-0055723P.
PR   18-AUG-1997;   97US-0055947P.
PR   18-AUG-1997;   97US-0055948P.
PR   18-AUG-1997;   97US-0055949P.
PR   18-AUG-1997;   97US-0055950P.
PR   18-AUG-1997;   97US-0055953P.
PR   18-AUG-1997;   97US-0055954P.
PR   18-AUG-1997;   97US-0055964P.
PR   18-AUG-1997;   97US-0055984P.
PR   18-AUG-1997;   97US-0056360P.
PR   12-SEP-1997;   97US-0058660P.
PR   12-SEP-1997;   97US-0058661P.
PR   12-SEP-1997;   97US-0058664P.
PR   12-SEP-1997;   97US-0058785P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Fischer CL,  Rosen CA,  Soppet DR,  Ruben SM,  Kyaw H,  Li Y,  Zeng Z;
PI   Lafleur DW,  Moore PA,  Shi Y,  Olsen HS,  Ebner R,  Brewer LA;
XX
DR   WPI; 1999-120770/10.
XX
PT   New isolated human genes and the secreted polypeptides they encode -
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders.
XX
PS   Disclosure; Page 101; 464pp; English.
XX
CC   This sequence represents a fragment of a secreted human protein encoded
CC   by the nucleic acid molecule detailed in the descriptor line. The gene
CC   can be used to generate fusion proteins by linking to the gene to a human
CC   immunoglobulin Fc portion (e.g. AAX27302) for increasing the stability of
CC   the fused protein as compared to the human protein only. The invention
CC   relates to 123 novel genes and their fragments (nucleic acid sequences:
CC   AAX27311-X27449; amino acid sequences AAY02650-Y02788) which are useful
CC   for preventing, treating or ameliorating medical conditions e.g. by
CC   protein or gene therapy. Also, pathological conditions can be diagnosed
CC   by determining the amount of the new polypeptides in a sample or by
CC   determining the presence of mutations in the new polynucleotides.
CC   Specific uses are described for each of the 123 polynucleotides, based on
CC   which tissues they are most highly expressed in (see AAX27311 for
CC   described uses)
XX
SQ   Sequence 30 AA;

  Query Match             100.0%;  Score 15;  DB 2;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 8.9e+02;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         12 RTLTF 16


RESULT 23
ADA07705
ID   ADA07705 standard; peptide; 30 AA.
XX
AC   ADA07705;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human secreted protein from gene 85 peptide #5.
XX
KW   Immunosuppressive; dermatological; antiinflammatory; antiallergic;
KW   antiarthritic; human; autoimmune disease; autoimmune disorder; lupus;
KW   transplant rejection; allergic reaction; arthritis;
KW   squamous cell E48 antigen.
XX
OS   Homo sapiens.
XX
PN   US2003064412-A1.
```

XX
PD 03-APR-2003.
XX
PF 30-OCT-2001; 2001US-00984490.
XX
PR 08-JUL-1997; 97US-0051916P.
PR 08-JUL-1997; 97US-0051918P.
PR 08-JUL-1997; 97US-0051919P.
PR 08-JUL-1997; 97US-0051920P.
PR 08-JUL-1997; 97US-0051925P.
PR 08-JUL-1997; 97US-0051926P.
PR 08-JUL-1997; 97US-0051928P.
PR 08-JUL-1997; 97US-0051929P.
PR 08-JUL-1997; 97US-0051930P.
PR 08-JUL-1997; 97US-0051931P.
PR 08-JUL-1997; 97US-0051932P.
PR 08-JUL-1997; 97US-0052732P.
PR 08-JUL-1997; 97US-0052733P.
PR 08-JUL-1997; 97US-0052793P.
PR 08-JUL-1997; 97US-0052795P.
PR 08-JUL-1997; 97US-0052803P.
PR 18-AUG-1997; 97US-0055684P.
PR 18-AUG-1997; 97US-0055722P.
PR 18-AUG-1997; 97US-0055723P.
PR 18-AUG-1997; 97US-0055947P.
PR 18-AUG-1997; 97US-0055948P.
PR 18-AUG-1997; 97US-0055949P.
PR 18-AUG-1997; 97US-0055950P.
PR 18-AUG-1997; 97US-0055953P.
PR 18-AUG-1997; 97US-0055954P.
PR 18-AUG-1997; 97US-0055964P.
PR 18-AUG-1997; 97US-0055984P.
PR 18-AUG-1997; 97US-0056360P.
PR 12-SEP-1997; 97US-0058660P.
PR 12-SEP-1997; 97US-0058661P.
PR 12-SEP-1997; 97US-0058664P.
PR 12-SEP-1997; 97US-0058785P.
PR 07-JUL-1998; 98WO-US013684.
PR 08-JAN-1999; 99US-00227357.
XX
PA (FISC/) FISCHER C L.
PA (ROSE/) ROSEN C A.
PA (SOPP/) SOPPET D R.
PA (RUBE/) RUBEN S M.
PA (KYAW/) KYAW H.
PA (LIYY/) LI Y.
PA (ZENG/) ZENG Z.
PA (LAFL/) LAFLEUR D W.
PA (MOOR/) MOORE P A.
PA (SHIY/) SHI Y.
PA (OLSE/) OLSEN H S.
PA (EBNE/) EBNER R.
PA (BREW/) BREWER L A.
XX
PI Fischer CL, Rosen CA, Soppet DR, Ruben SM, Kyaw H, Li Y, Zeng Z;
PI Lafleur DW, Moore PA, Shi Y, Olsen HS, Ebner R, Brewer LA;
XX
DR WPI; 2003-540785/51.
XX
PT Novel antibody which specifically binds to a secreted protein useful for
PT diagnosing and treating lupus, arthritis, allergic reactions, arthritis.
XX
PS Disclosure; Page 57; 355pp; English.
XX
CC The invention relates to an isolated antibody or its portion that
CC specifically binds to a protein that shares sequence homology with human
CC squamous cell E48 antigen, and consists of amino acid residues 21-116 or
CC 1-116 the protein appearing as ADA07417 (one of 123 disclosed novel human
CC secreted proteins encoded by 123 novel genes), or a protein consisting of
CC amino acid sequence of secreted or full-length polypeptide encoded by
CC HLHFPO3 cDNA contained in ATCC Deposit No. 209126. The antibody is
CC produced by immunising an animal with amino acid residues 21-116 of
CC ADA07417, or with a protein consisting of amino acid sequence of the
CC secreted polypeptide encoded by the HLHFPO3 cDNA contained in ATCC
CC Deposit No. 209126, respectively. Also included are an isolated cell that
CC produces the antibody and a hybridoma that produces the antibody. The
CC antibody is a monoclonal, polyclonal, chimeric, humanised or human
CC antibody. Optionally, the antibody is a Fab fragment, and is labelled by
CC a label chosen from enzyme label, a radioisotope, and a fluorescent
CC label. The antibody is useful as a probe for differential identification
CC of tissues or cell types in which ADA07417 is expressed. The antibody is
CC also for diagnosis and treatment of autoimmune diseases and disorders,
CC such as lupus, transplant rejection, allergic reactions, and arthritis.
CC The present sequence is a peptide/protein derived from one of the 123
CC novel secreted proteins.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 15; DB 7; Length 30;
Best Local Similarity 60.0%; Pred. No. 8.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 12 RTLTF 16

RESULT 24
ADN41296
ID ADN41296 standard; peptide; 30 AA.
XX
AC ADN41296;
XX
DT 17-JUN-2004 (first entry)
XX
DE Novel human secreted protein fragment seqid 418.
XX
KW immunomodulator; immunosuppressive; antiinflammatory; dermatological;
KW antiarthritic; antirheumatic; neuroprotective; antianaemic; muscular;
KW antiallergic; antiasthmatic; gastrointestinal; anticoagulant;
KW thrombolytic; antiarteriosclerotic; cardiant; cytostatic; nephrotropic;
KW cardiovascular; respiratory; gene therapy; secreted protein;
KW chromosome identification; hybrid mapping; gene expression control;
KW immune system disorder; immunodeficiency; Chediak-Higashi syndrome;
KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
KW multiple sclerosis; haemolytic anaemia; myasthenia gravis;
KW allergic reaction; asthma; inflammatory condition;
KW inflammatory bowel disease; B cell stimulator; T cell activator;
KW blood-related disorder; eosinophilia; thrombosis; thromboembolism;
KW atherosclerosis; myocardial infarction; angina; anaemia;
KW hyperproliferative disorder; cancer; renal disorder;
KW chronic kidney failure; renal tubular acidosis; kidney stone;
KW cardiovascular disorder; respiratory disorder; human.
XX
OS Homo sapiens.
XX
PN US2004044191-A1.
XX
PD 04-MAR-2004.
XX
PF 10-OCT-2001; 2001US-00973278.
XX
PR 08-JUL-1997; 97US-0051916P.
PR 08-JUL-1997; 97US-0051918P.
PR 08-JUL-1997; 97US-0051919P.
PR 08-JUL-1997; 97US-0051920P.
PR 08-JUL-1997; 97US-0051925P.
PR 08-JUL-1997; 97US-0051926P.
PR 08-JUL-1997; 97US-0051928P.
PR 08-JUL-1997; 97US-0051929P.
PR 08-JUL-1997; 97US-0051930P.
PR 08-JUL-1997; 97US-0051931P.
PR 08-JUL-1997; 97US-0051932P.
PR 08-JUL-1997; 97US-0052732P.
PR 08-JUL-1997; 97US-0052733P.
PR 08-JUL-1997; 97US-0052793P.

PR 08-JUL-1997; 97US-0052795P.
PR 08-JUL-1997; 97US-0052803P.
PR 18-AUG-1997; 97US-0055684P.
PR 18-AUG-1997; 97US-0055722P.
PR 18-AUG-1997; 97US-0055723P.
PR 18-AUG-1997; 97US-0055947P.
PR 18-AUG-1997; 97US-0055948P.
PR 18-AUG-1997; 97US-0055949P.
PR 18-AUG-1997; 97US-0055950P.
PR 18-AUG-1997; 97US-0055953P.
PR 18-AUG-1997; 97US-0055954P.
PR 18-AUG-1997; 97US-0055964P.
PR 18-AUG-1997; 97US-0055984P.
PR 18-AUG-1997; 97US-0056360P.
PR 12-SEP-1997; 97US-0058660P.
PR 12-SEP-1997; 97US-0058661P.
PR 12-SEP-1997; 97US-0058664P.
PR 12-SEP-1997; 97US-0058785P.
PR 07-JUL-1998; 98WO-US013684.
PR 08-JAN-1999; 99US-00227357.
PR 13-OCT-2000; 2000US-0239899P.
XX
PA (FISC/) FISCHER C L.
PA (ROSE/) ROSEN C A.
PA (SOPP/) SOPPET D R.
PA (RUBE/) RUBEN S M.
PA (KYAW/) KYAW H.
PA (LIYY/) LI Y.
PA (ZENG/) ZENG Z.
PA (LAFL/) LAFLEUR D W.
PA (MOOR/) MOORE P A.
PA (SHIY/) SHI Y.
PA (OLSE/) OLSEN H.
PA (EBNE/) EBNER R.
PA (BIRS/) BIRSE C E.
XX
PI Fischer CL, Rosen CA, Soppet DR, Ruben SM, Kyaw H, Li Y, Zeng Z;
PI Lafleur DW, Moore PA, Shi Y, Olsen H, Ebner R, Birse CE;
XX
DR WPI; 2004-225733/21.
XX
PT New isolated nucleic acid encoding human proteins, useful for treating,
PT preventing or diagnosing e.g. rheumatoid arthritis, multiple sclerosis,
PT anemia, inflammatory bowel disease, atherosclerosis, cancers, chronic
PT kidney failure.
XX
PS Disclosure; SEQ ID NO 418; 372pp; English.
XX
CC The invention describes novel human secreted proteins and the nucleotides
CC encoding them. The polynucleotides are useful in chromosome
CC identification, for radiation hybrid mapping, in controlling gene
CC expression, in gene therapy or as molecular weight markers. The
CC polynucleotides and polypeptides are useful for diagnosing, treating or
CC preventing diseases of the immune system, immunodeficiencies, e.g.
CC Chediak-Higashi syndrome, autoimmune diseases, e.g. systemic lupus
CC erythematosus, rheumatoid arthritis, multiple sclerosis, haemolytic
CC anaemia or myasthenia gravis, allergic reactions, e.g. asthma,
CC inflammatory conditions, e.g. inflammatory bowel disease. They can also
CC be used as a stimulator of B cell responsiveness to pathogens or as an
CC activator of T cells. The polynucleotides and polypeptides are also
CC useful for treating or preventing blood-related disorders, e.g.
CC eosinophilia, thrombosis, thromboembolism, atherosclerosis, myocardial
CC infarction, unstable angina or anaemia. They can also be used for
CC treating, preventing or diagnosing hyperproliferative disorders
CC (cancers), renal disorders (chronic kidney failure, renal tubular
CC acidosis or kidney stones), cardiovascular disorders or respiratory
CC disorders. This is the amino acid sequence of a novel human secreted
CC protein fragment. Note: This sequence is available in electronic format
CC from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20040044191.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 15; DB 8; Length 30;
Best Local Similarity 60.0%; Pred. No. 8.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 12 RTLTF 16

RESULT 25
AAW45039
ID AAW45039 standard; peptide; 31 AA.
XX
AC AAW45039;
XX
DT 27-APR-1998 (first entry)
XX
DE Immunomodulatory peptide D22184AA.
XX
KW Immunomodulator; immunosuppressant; immunostimulator; treatment;
KW transplant rejection; autoimmune disease; cancer; infection.
XX
OS Synthetic.
XX
PN WO9739023-A1.
XX
PD 23-OCT-1997.
XX
PF 04-APR-1997; 97WO-SE000574.
XX
PR 12-APR-1996; 96SE-00001422.
PR 23-SEP-1996; 96SE-00003469.
XX
PA (ASTR) ASTRA AB.
XX
PI Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;
XX
DR WPI; 1997-526397/48.
DR N-PSDB; AAV05459.
XX
PT Nucleic acids encoding cysteine- or methionine-containing peptide(s)
PT which have immuno:stimulatory or immunosuppressive activity - can be used
PT to treat, e.g. cancers, infection, auto:immune disease or transplant
PT rejection.
XX
PS Claim 22; Page 162; 183pp; English.
XX
CC The present peptide is an immunosuppressant or immunostimulator. An
CC immunosuppressant can be used to treat transplant rejection or autoimmune
CC disease, e.g. rheumatoid arthritis, systemic lupus erythematosis,
CC Sjogren's syndrome, scleroderma, mixed connective tissue disease,
CC dermatomyositis, polymyositis, Reiter's syndrome, Behcet's disease, type
CC I diabetes, Hashimoto's thyroiditis, Graves' disease, multiple sclerosis,
CC myasthenia gravis, encephalomyelitis, phemphigus vulgaris, vegetans or
CC foliaceus, Senear-Usher syndrome or Brazilian phemphigus. An
CC immunostimulator can be used to treat conditions such as cancer or
CC infection
XX
SQ Sequence 31 AA;

Query Match 100.0%; Score 15; DB 2; Length 31;
Best Local Similarity 60.0%; Pred. No. 9.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 23 RALAF 27

RESULT 26
AAY09486
ID AAY09486 standard; peptide; 31 AA.
XX

AC AAY09486;
XX
DT 14-JUL-1999 (first entry)
XX
DE Immunoactive peptide containing a mammalian signal peptide #1.
XX
KW Immunoactive; immunomodulation; immunosuppression; immunostimulation;
KW immune response; immunoreactive; autoimmune disease.
XX
OS Synthetic.
XX
PN WO9919347-A1.
XX
PD 22-APR-1999.
XX
PF 06-OCT-1998; 98WO-SE001801.
XX
PR 10-OCT-1997; 97US-00949024.
XX
PA (ASTR) ASTRA AB.
XX
PI Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;
XX
DR WPI; 1999-287953/24.
XX
PT Synthetic genes encoding immunoreactive peptides containing cysteine or
PT methionine.
XX
PS Disclosure; Page 12; 104pp; English.
XX
CC The present invention describes nucleic acid molecules comprising a
CC coding sequence encoding an immunoreactive peptide and further encoding a
CC protein targeting sequence. The nucleic acid is administered to a patient
CC so that its expression product, an immunoactive peptide, modulates an
CC immune response in a patient. The nucleic acid can also be used to treat
CC cancer, either after surgery to remove a portion of the cancer or after
CC ionizing radiation. A cytokine is also administered in conjunction with
CC the nucleic acid. Cells containing the nucleic acid molecule can also be
CC used for treatment. The immunoactive peptide is immunosuppressive and can
CC be used in patients with autoimmune disease. The present sequence
CC represents an immunoactive peptide from the present invention
XX
SQ Sequence 31 AA;

Query Match 100.0%; Score 15; DB 2; Length 31;
Best Local Similarity 60.0%; Pred. No. 9.2e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 23 RALAF 27

RESULT 27
ADV99762
ID ADV99762 standard; peptide; 32 AA.
XX
AC ADV99762;
XX
DT 07-APR-2005 (first entry)
XX
DE Glucanase polypeptide signal sequence #32.
XX
KW Glucanase; hydrolysis; enzyme engineering; insect resistance;
KW antibacterial; enzyme.
XX
OS Unidentified.
XX
PN WO2005003319-A2.
XX
PD 13-JAN-2005.
XX
PF 02-JUL-2004; 2004WO-US021492.
XX
PR 02-JUL-2003; 2003US-0484725P.
XX
PA (DIVE-) DIVERSA CORP.
XX
PI Steer B, Callen W, Healey S, Pulliam D;
XX
DR WPI; 2005-112400/12.
XX
PT New isolated recombinant glucanase polypeptide, useful in e.g., pulp
PT treatment, food processing, animal feeds, preparing dough, preparing fuel
PT products, and in brewing.
XX
PS Disclosure; Page 171; 289pp; English.
XX
CC The invention relates to a glucanase polypeptide. The invention also
CC relates to a nucleic acid encoding the polypeptide. The polypeptide is
CC useful for hydrolyzing a cellulose (derivative) or a hemicellulose, e.g.,
CC in a wood, paper pulp or a paper product, and for catalyzing hydrolysis
CC of glucan in a feed, food product or a beverage. The feed, food product
CC or beverage comprises a cereal-based animal feed, dough, wort or beer, or
CC fruit or vegetable. The polypeptide is also useful for catalyzing
CC hydrolysis of xylans in a microbe, plant cell, insect, yeast or mammalian
CC cell, for catalyzing hydrolysis of a sugar to make a fuel product, for
CC identifying its binding agents (e.g. modulators), and as a nutritional
CC supplement in an animal diet, which involves preparing a nutritional
CC supplement containing a glucanase enzyme comprising at least thirty
CC contiguous amino acids of the glucanase polypeptide, and administering
CC the nutritional supplement to an animal to increase utilization of xylan
CC contained in feed or food ingested by the animal, where the animal is a
CC human, ruminant or a monogastric animal. The glucanase polypeptide is
CC further useful for reducing lignin in a paper, wood or wood product, for
CC eliminating or protecting animals from glucan containing microorganisms
CC (especially Salmonella), in oral care products, for producing or
CC improving the flavor or texture of dairy products, and for producing
CC small molecules. This sequence represents a glucanase polypeptide signal
CC sequence used in the scope of the invention.
XX
SQ Sequence 32 AA;

Query Match 100.0%; Score 15; DB 9; Length 32;
Best Local Similarity 60.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 8 RTLSF 12

RESULT 28
AAB44770
ID AAB44770 standard; protein; 33 AA.
XX
AC AAB44770;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 9 SEQ ID NO:69.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; skin aging; food additive; preservative.
XX
OS Homo sapiens.
XX
PN WO200058336-A1.
XX
PD 05-OCT-2000.
XX

PF 23-MAR-2000; 2000WO-US007726.
XX
PR 26-MAR-1999; 99US-0126597P.
PR 07-JAN-2000; 2000US-0174877P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-602355/57.
DR N-PSDB; AAC79807.
XX
PT Nucleic acid encoding human secreted proteins, used to treat, prevent,
PT ameliorate or diagnose medical conditions such as cancer, and autoimmune
PT diseases.
XX
PS Claim 11; Page 359; 391pp; English.
XX
CC The polynucleotide sequences given in AAC79799 to AAC79848 encode the
CC human secreted proteins given in AAB44762 to AAB44811. AAB44812 to
CC AAB44829 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are used in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities are:
CC immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
CC cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and polypeptides can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders, cardiovascular
CC disorders, cerebrovascular disorders, angiogenesis, nervous system
CC disorders, infections caused by bacteria, viruses and fungi and ocular
CC disorders. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. AAC79790 to AAC79798 and
CC AAB44761 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 15; DB 3; Length 33;
Best Local Similarity 60.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 6 RSLSF 10

RESULT 29
AAM32187
ID AAM32187 standard; protein; 33 AA.
XX
AC AAM32187;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #6224 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 32456; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 15; DB 4; Length 33;
Best Local Similarity 60.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RTLSF 18

RESULT 30
AAM60971
ID AAM60971 standard; protein; 33 AA.
XX
AC AAM60971;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33076.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 33076; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 15; DB 4; Length 33;
Best Local Similarity 60.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RTLSF 18

RESULT 31
ABG55401
ID ABG55401 standard; peptide; 33 AA.
XX
AC ABG55401;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID No 34049.
XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 34049; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 15; DB 4; Length 33;
Best Local Similarity 60.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RTLSF 18

RESULT 32
ABG53589
ID ABG53589 standard; peptide; 33 AA.
XX
AC ABG53589;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID No 32237.
XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 32237; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 15; DB 4; Length 33;
Best Local Similarity 60.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RTLSF 18

RESULT 33
ABG43538
ID ABG43538 standard; peptide; 33 AA.
XX
AC ABG43538;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 33203.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 33203; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 15; DB 5; Length 33;
Best Local Similarity 60.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RTLSF 18

RESULT 34
ABG41720
ID ABG41720 standard; peptide; 33 AA.
XX
AC ABG41720;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 31385.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 31385; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 15; DB 5; Length 33;
Best Local Similarity 60.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 14 RTLSF 18

RESULT 35
AAB44347
ID AAB44347 standard; protein; 35 AA.
XX
AC AAB44347;
XX
DT 14-FEB-2001 (first entry)
XX
DE Human secreted protein encoded by gene 13 clone HSREC72.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200058358-A1.
XX
PD 05-OCT-2000.
XX
PF 23-MAR-2000; 2000WO-US007725.
XX
PR 26-MAR-1999; 99US-0126602P.
PR 14-JAN-2000; 2000US-0176063P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-594640/56.
DR N-PSDB; AAC79009.
XX
PT Fourty nine nucleic acid molecules encoding human secreted proteins,
PT useful in the prevention, treatment and diagnosis of cancer, immune
PT disorders, cardiovascular disorders and neurological diseases.
XX
PS Claim 11; Page 342; 367pp; English.
XX
CC Sequences AAB44335-B44382 represent the amino acid sequences of 49 human
CC secreted proteins encoded by the genes AAC69084-C69119. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections.
XX
SQ Sequence 35 AA;

Query Match 100.0%; Score 15; DB 3; Length 35;
Best Local Similarity 60.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 9 RSLTF 13

RESULT 36
ABO56517
ID ABO56517 standard; protein; 35 AA.
XX
AC ABO56517;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #2751.
XX
KW Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.

```
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 45; SEQ ID NO 30151; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe protein of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 35 AA;

  Query Match           100.0%;  Score 15;  DB 8;  Length 35;
  Best Local Similarity   60.0%;  Pred. No. 1e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         18 RALAF 22


RESULT 37
AAB63066
ID   AAB63066 standard; protein; 36 AA.
XX
AC   AAB63066;
XX
DT   26-MAR-2001  (first entry)
XX
DE   Human secreted protein sequence encoded by gene 18 SEQ ID NO:76.
XX
KW   Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; vulnerary; gene therapy; neoplasm;
KW   autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW   cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW   cerebral ischaemia; angiogenesis; nervous system disorder; infection;
KW   Alzheimer's disease; ocular disorder; corneal infection; wound healing;
KW   skin aging; food additive; preservative.
XX
OS   Homo sapiens.
XX
PN   WO200061748-A1.
XX
PD   19-OCT-2000.
XX
PF   06-APR-2000; 2000WO-US008982.
XX
PR   09-APR-1999;   99US-0128696P.
PR   14-JAN-2000; 2000US-0176069P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-638566/61.
DR   N-PSDB; AAF22333.
XX
PT   New nucleic acid molecules encoding 48 human secreted proteins for
PT   diagnosing, preventing, treating or ameliorating medical conditions and
PT   used as food additives or preservatives.
XX
PS   Claim 11; Page 436; 480pp; English.
XX
CC   AAF22316 to AAF22363 encode the human secreted proteins given in AAB63049
CC   to AAB63096. AAB63097 to AAB63132 represent more human secreted proteins
CC   and polypeptides homologous to them. Human secreted proteins have
CC   activities based on the tissues and cells the genes are expressed in.
CC   Examples of activities include: immunosuppressive; antiarthritic;
CC   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
CC   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
CC   fungicide; ophthalmological; and vulnerary. The polynucleotides and
CC   proteins can be used to prevent, treat or ameliorate a medical condition
CC   in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or
CC   sheep. They are also used in diagnosing a pathological condition or
CC   susceptibility to a pathological condition. Disorders which are diagnosed
CC   or treated include autoimmune diseases e.g. rheumatoid arthritis,
CC   hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC   cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC   e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC   Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC   ocular disorders e.g. corneal infection. The polypeptides can also be
CC   used to aid wound healing and epithelial cell proliferation, to prevent
CC   skin aging due to sunburn, to maintain organs before transplantation, for
CC   supporting cell culture of primary tissues, to regenerate tissues and in
CC   chemotaxis. The polypeptides can also be used as a food additive or
CC   preservative to increase or decrease storage capabilities, fat content,
CC   lipid, protein, carbohydrate, vitamins, minerals, cofactors and other
CC   nutritional components. AAF22307 to AAF22315 and AAB63048 represent
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 36 AA;

  Query Match           100.0%;  Score 15;  DB 3;  Length 36;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
```

```
Db         31 RSLAF 35


RESULT 38
ABG68830
ID   ABG68830 standard; protein; 38 AA.
XX
AC   ABG68830;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Cytochrome P450 3A7 (CYP3A7) variant C-terminal ORF #2.
XX
KW   Cytochrome P450; CYP3AP1; CYP3AP2; CYP3A43; CYP3A4; CYP3A5; CYP3A7;
KW   drug metabolism; drug design; drug screening.
XX
OS   Synthetic.
XX
PN   WO200244213-A1.
XX
PD   06-JUN-2002.
XX
PF   28-NOV-2001; 2001WO-SE002631.
XX
PR   28-NOV-2000; 2000SE-00004366.
PR   11-JUN-2001; 2001SE-00002061.
XX
PA   (ZAPH/) ZAPHIROPOULOS P G.
PA   (FINT/) FINTA C.
XX
PI   Zaphiropoulos PG,  Finta C;
XX
DR   WPI; 2002-557532/59.
XX
PT   Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
PT   of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
PT   metabolism, in drug design and drug screening.
XX
PS   Disclosure; Fig 2; 131pp; English.
XX
CC   The invention describes a cytochrome P450 protein (I) in which CYP3A43
CC   exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
CC   fragments, variants and multiples of (I) having essentially the same
CC   characteristics. (I) Is useful as a medicament, and for evaluating drug
CC   metabolism, in drug design, and drug screening, and in tests for
CC   adjusting the dose of drugs. This is the amino acid sequence of a
CC   cytochrome P450 3A7 (CYP3A7) variant C-terminal
XX
SQ   Sequence 38 AA;

  Query Match             100.0%;  Score 15;  DB 5;  Length 38;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         15 RSLAF 19


RESULT 39
AAG15201
ID   AAG15201 standard; protein; 40 AA.
XX
AC   AAG15201;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 15360.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
```

PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 15; DB 3; Length 40;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 RXLXF 8
              | | |
Db         12 RSLSF 16
```

RESULT 40
ADD90505
ID ADD90505 standard; protein; 40 AA.
XX
AC ADD90505;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel human secreted protein seq id 68 protein feature seq id 320.
XX
KW gene therapy; cytostatic; cancer; human; secreted protein.
XX
OS Homo sapiens.
XX
PN US2003199683-A1.
XX
PD 23-OCT-2003.
XX
PF 30-MAR-2001; 2001US-00820649.
XX
PR 30-JUL-1997; 97US-0054209P.

```
PR   30-JUL-1997;   97US-0054211P.
PR   30-JUL-1997;   97US-0054212P.
PR   30-JUL-1997;   97US-0054213P.
PR   30-JUL-1997;   97US-0054214P.
PR   30-JUL-1997;   97US-0054215P.
PR   30-JUL-1997;   97US-0054217P.
PR   30-JUL-1997;   97US-0054218P.
PR   30-JUL-1997;   97US-0054234P.
PR   30-JUL-1997;   97US-0054236P.
PR   18-AUG-1997;   97US-0055968P.
PR   18-AUG-1997;   97US-0055969P.
PR   18-AUG-1997;   97US-0055972P.
PR   19-AUG-1997;   97US-0056534P.
PR   19-AUG-1997;   97US-0056543P.
PR   19-AUG-1997;   97US-0056554P.
PR   19-AUG-1997;   97US-0056561P.
PR   19-AUG-1997;   97US-0056727P.
PR   19-AUG-1997;   97US-0056729P.
PR   19-AUG-1997;   97US-0056730P.
PR   29-JUL-1998;   98WO-US015949.
PR   26-JAN-1999;   99US-00236557.
PR   21-SEP-2000; 2000US-00666987.
XX
PA   (RUBE/) RUBEN S M.
PA   (FENG/) FENG P.
PA   (LAFL/) LAFLEUR D W.
PA   (MOOR/) MOORE P A.
PA   (SHIY/) SHI Y.
PA   (KYAW/) KYAW H.
PA   (LIYY/) LI Y.
PA   (ZENG/) ZENG Z.
PA   (CART/) CARTER K C.
PA   (ENDR/) ENDRESS G A.
PA   (WEIY/) WEI Y.
PA   (FANP/) FAN P.
PA   (ROSE/) ROSEN C A.
XX
PI   Ruben SM,  Feng P,  Lafleur DW,  Moore PA,  Shi Y,  Kyaw H,  Li Y;
PI   Zeng Z,  Carter KC,  Endress GA,  Wei Y,  Fan P,  Rosen CA;
XX
DR   WPI; 2003-852813/79.
XX
PT   New nucleic acid molecule, useful for preparing a medicament for
PT   preventing, treating or ameliorating a medical condition e.g., cancer.
XX
PS   Disclosure; SEQ ID NO 320; 213pp; English.
XX
CC   The invention describes novel isolated human nucleic acids. The nucleic
CC   acid is useful for preparing a medicament for preventing, treating or
CC   ameliorating a medical condition e.g., cancer, and in gene therapy. This
CC   is the amino acid sequence of polypeptide feature of a novel human
CC   secreted protein of the invention.
XX
SQ   Sequence 40 AA;

  Query Match          100.0%;  Score 15;  DB 7;  Length 40;
  Best Local Similarity   60.0%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          2 RSLSF 6


RESULT 41
ADG90324
ID   ADG90324 standard; protein; 40 AA.
XX
AC   ADG90324;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human secreted protein gene 58 extra polypeptide #2.
XX
KW   Secreted protein; gene therapy; neural disorder; immune system disorders;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   proliferative disorder; cancer; systemic lupus erythematosus;
KW   rheumatoid arthritis; multiple sclerosis; thyroiditis; anaemia;
KW   Grave's disease; diabetes; hepatitis; asthma; allergy; nephritis;
KW   Parkinson's disease; Alzheimer's disease; atherosclerosis;
KW   myocardial infarction; AIDS; infection; human.
XX
OS   Homo sapiens.
XX
PN   US2003166541-A1.
XX
PD   04-SEP-2003.
XX
PF   04-JUN-2002; 2002US-00160162.
XX
PR   30-JUL-1997;   97US-0054209P.
PR   30-JUL-1997;   97US-0054211P.
PR   30-JUL-1997;   97US-0054212P.
PR   30-JUL-1997;   97US-0054213P.
PR   30-JUL-1997;   97US-0054214P.
PR   30-JUL-1997;   97US-0054215P.
PR   30-JUL-1997;   97US-0054217P.
PR   30-JUL-1997;   97US-0054218P.
PR   30-JUL-1997;   97US-0054234P.
PR   30-JUL-1997;   97US-0054236P.
PR   18-AUG-1997;   97US-0055968P.
PR   18-AUG-1997;   97US-0055969P.
PR   18-AUG-1997;   97US-0055972P.
PR   19-AUG-1997;   97US-0056534P.
PR   19-AUG-1997;   97US-0056543P.
PR   19-AUG-1997;   97US-0056554P.
PR   19-AUG-1997;   97US-0056561P.
PR   19-AUG-1997;   97US-0056727P.
PR   19-AUG-1997;   97US-0056729P.
PR   19-AUG-1997;   97US-0056730P.
PR   29-JUL-1998;   98WO-US015949.
PR   26-JAN-1999;   99US-00236557.
PR   05-JUN-2001; 2001US-0295558P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Feng P,  Lafleur DW,  Moore PA,  Shi Y,  Kyaw H,  Li Y;
PI   Zeng Z,  Carter KC,  Endress GA,  Wei Y,  Fan P,  Rosen CA;
XX
DR   WPI; 2003-874923/81.
XX
PT   Nucleic acids encoding 83 secreted polypeptides, useful for preventing,
PT   diagnosing and treating disorders related to their aberrant expression
PT   and activity.
XX
PS   Disclosure; SEQ ID NO 320; 308pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule encoding a
CC   secreted protein that is at least 95% identical to a polynucleotide
CC   fragment of any of the nucleotide sequences listed in table 1A of the
CC   specification, which is hybridisable to the nucleotide sequences, a
CC   polynucleotide encoding a polypeptide (or a polypeptide fragment, domain
CC   or epitope of any of the amino acid sequences) listed in table 1A of the
CC   specification, a polynucleotide which is an (allelic) variant of the
CC   nucleotide sequences listed in the specification, a polynucleotide which
CC   encodes a species homologue of the above amino acid sequences, a
CC   polynucleotide capable of hybridising under stringent conditions to any
CC   of the above polynucleotides, where the polynucleotide does not hybridise
CC   under stringent conditions to a nucleic acid molecule having a nucleotide
CC   sequence of only A or T residues. Also included are a recombinant vector
CC   comprising the above nucleic acid molecule, making a recombinant host
CC   cell comprising the above nucleic acid molecule, an isolated polypeptide
CC   comprising a sequence that is at least 95% identical to the polypeptide
CC   (or its fragment, domain, epitope, secreted form, (allelic) variant or
CC   homologue) encoded by the above nucleic acid molecule, an isolated
```

CC antibody that binds specifically to the above polypeptide, a recombinant
CC host cell produced by the above method and that expresses the above
CC polypeptide, making an isolated polypeptide, preventing, treating or
CC ameliorating a medical condition, diagnosing a pathological condition or
CC a susceptibility to a pathological condition in a subject, identifying a
CC binding partner to the above polypeptide, the gene corresponding to the
CC cDNA sequence given in the specification, and identifying an activity in
CC a biological assay. The nucleic acid molecule and polypeptide are useful
CC in diagnosing, preventing, prognosing or treating diseases or disorders
CC associated with aberrant expression and/or activity of the above
CC polypeptide, such as neural disorders, immune system disorders, muscular
CC disorders, reproductive disorders, gastrointestinal disorders, pulmonary
CC disorders, cardiovascular disorders, renal disorders, proliferative
CC disorders and/or cancers. In particular, these diseases are systemic
CC lupus erythematosus, rheumatoid arthritis, multiple sclerosis,
CC thyroiditis, anaemia, Grave's disease, diabetes, hepatitis, asthma,
CC allergies, nephritis, Parkinson's disease, Alzheimer's disease,
CC atherosclerosis, myocardial infarction, AIDS and infections. The methods
CC may be used for identifying agonists and antagonists of the
CC polynucleotide and polypeptide. The present sequence is a protein from
CC one of the 83 disclosed secreted protein genes.
XX
SQ Sequence 40 AA;

Query Match 100.0%; Score 15; DB 7; Length 40;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 2 RSLSF 6

RESULT 42
ABO57531
ID ABO57531 standard; protein; 40 AA.
XX
AC ABO57531;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #3765.
XX
KW Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 45; SEQ ID NO 31165; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 40 AA;

Query Match 100.0%; Score 15; DB 8; Length 40;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 22 RSLSF 26

RESULT 43
ADY25664
ID ADY25664 standard; peptide; 40 AA.
XX
AC ADY25664;
XX
DT 05-MAY-2005 (first entry)
XX
DE Novel human secreted protein-related peptide SeqID320.
XX
KW cancer; cytostatic; arthritis; antiarthritic; asthma; antiasthmatic;
KW acquired immune deficiency syndrome; rheumatoid arthritis; antirheumatic;
KW inflammatory bowel disease; antiinflammatory; gastrointestinal-gen.;
KW sepsis; antibacterial; immunosuppressive; acne; antiseborrheic;
KW dermatological; psoriasis; antipsoriatic; atherosclerosis;
KW antiarteriosclerotic; cerebrovascular ischemia; cerebroprotective;
KW vasotropic; thrombosis; wound healing; vulnerary; Alzheimers disease;
KW neuroprotective; nootropic; parkinsons disease; antiparkinsonian; autism;
KW obsessive-compulsive disorder; tranquilizer; graft versus host disease;
KW immune disorder; hematological disease; inflammation; infection;
KW hyperproliferative disorders; renal disease; nephrotropic;
KW cardiovascular disease; cardiovascular-gen.; respiratory disorder;
KW neurological disease; neuroprotective; endocrine disease;
KW reproductive disorders (general); gynecological.
XX
OS Homo sapiens.
XX
PN US2005037467-A1.

```
XX
PD   17-FEB-2005.
XX
PF   09-SEP-2004; 2004US-00936773.
XX
PR   30-JUL-1997;   97US-0054209P.
PR   30-JUL-1997;   97US-0054211P.
PR   30-JUL-1997;   97US-0054212P.
PR   30-JUL-1997;   97US-0054213P.
PR   30-JUL-1997;   97US-0054214P.
PR   30-JUL-1997;   97US-0054215P.
PR   30-JUL-1997;   97US-0054217P.
PR   30-JUL-1997;   97US-0054218P.
PR   30-JUL-1997;   97US-0054234P.
PR   30-JUL-1997;   97US-0054236P.
PR   18-AUG-1997;   97US-0055968P.
PR   18-AUG-1997;   97US-0055969P.
PR   18-AUG-1997;   97US-0055972P.
PR   19-AUG-1997;   97US-0056534P.
PR   19-AUG-1997;   97US-0056543P.
PR   19-AUG-1997;   97US-0056554P.
PR   19-AUG-1997;   97US-0056561P.
PR   19-AUG-1997;   97US-0056727P.
PR   19-AUG-1997;   97US-0056729P.
PR   19-AUG-1997;   97US-0056730P.
PR   29-JUL-1998;   98WO-US015949.
PR   26-JAN-1999;   99US-00236557.
PR   21-SEP-2000; 2000US-00666984.
PR   30-MAR-2001; 2001US-00820649.
PR   05-JUN-2001; 2001US-0295558P.
PR   04-JUN-2002; 2002US-00160162.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Feng P,  Lafleur DW,  Moore PA,  Shi Y,  Kyaw H,  Li Y;
PI   Zeng Z,  Carter KC,  Endress GA,  Wei Y,  Fan P,  Rosen CA;
XX
DR   WPI; 2005-161941/17.
XX
PT   New nucleic acid molecule encoding human secreted protein, useful for
PT   preventing, treating, or ameliorating immune system, blood, inflammatory,
PT   infectious, cardiovascular, respiratory, neurological, endocrine, or
PT   reproductive disorders.
XX
PS   Disclosure; SEQ ID NO 320; 309pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of novel
CC   human secreted proteins and peptides. The DNA and protein sequences of
CC   the invention are useful for treating or ameliorating: cancer, arthritis,
CC   asthma, AIDS, rheumatoid arthritis, inflammatory bowel disease, sepsis,
CC   acne, psoriasis, atherosclerosis, stroke, thrombosis, wound healing,
CC   Alzheimer's Disease, Parkinson's Disease, autism, obsessive compulsive
CC   disorder, graft-versus-host diseases, immune system disorders, blood
CC   disorders, inflammatory conditions, infectious diseases,
CC   hyperproliferative disorders, renal disorders, cardiovascular disorders,
CC   respiratory disorders, neurological disorders, endocrine disorders, or
CC   reproductive disorders. The present sequence is that of a human peptide
CC   which is related to the novel human secreted proteins/peptides of the
CC   invention.
XX
SQ   Sequence 40 AA;

  Query Match             100.0%;  Score 15;  DB 9;  Length 40;
  Best Local Similarity   60.0%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db          2 RSLSF 6


RESULT 44
AAY25813
ID   AAY25813 standard; protein; 43 AA.
XX
AC   AAY25813;
XX
DT   04-OCT-1999  (first entry)
XX
DE   Human secreted protein fragment encoded from gene 23.
XX
KW   Secreted protein; human; treatment; diagnosis; therapy; cancer; tumour;
KW   neurodegenerative disorder; developmental abnormality; blood disorder;
KW   fetal deficiency; blood disorder; leukemia; immune system; inflammation;
KW   autoimmune disease; hepatic disease; renal disease; allergy; restenosis;
KW   ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;
KW   cardiovascular disorder; wound healing; stroke; arthritis; obesity;
KW   asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;
KW   metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   WO9938881-A1.
XX
PD   05-AUG-1999.
XX
PF   27-JAN-1999;   99WO-US001621.
XX
PR   30-JAN-1998;   98US-0073159P.
PR   30-JAN-1998;   98US-0073160P.
PR   30-JAN-1998;   98US-0073161P.
PR   30-JAN-1998;   98US-0073162P.
PR   30-JAN-1998;   98US-0073164P.
PR   30-JAN-1998;   98US-0073165P.
PR   30-JAN-1998;   98US-0073167P.
PR   30-JAN-1998;   98US-0073170P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ferrie AM,  Rosen CA,  Florence KA,  Carter KC,  Soppet DR;
PI   Yu G,  Florence C,  Young P,  Ni J,  Feng P,  Endress GA,  Janat F;
XX
DR   WPI; 1999-469315/39.
DR   N-PSDB; AAZ00432.
XX
PT   New isolated human genes and the secreted polypeptides they encode useful
PT   in, e.g. treatment of Alzheimer's.
XX
PS   Disclosure; Page 355; 393pp; English.
XX
CC   This invention describes novel human genes (see AAZ00410-Z00477) and the
CC   secreted proteins (see AAY25711-Y25778) and fragments (see AAY25779-
CC   Y25907) they encode. The polynucleotides and their corresponding secreted
CC   polypeptides are useful for preventing, treating or ameliorating medical
CC   conditions e.g. by protein or gene therapy. Also pathological conditions
CC   can be diagnosed by determining the amount of the new polypeptides in a
CC   sample or by determining the presence of mutations in the new
CC   polynucleotides. Specific uses are described for each of the 67
CC   polynucleotides of the invention, based on which tissues they are most
CC   highly expressed in, and include developing products for the diagnosis or
CC   treatment of cancer, tumours, neurodegenerative disorders, developmental
CC   abnormalities and fetal deficiencies, blood disorders, leukemias,
CC   diseases of the immune system, autoimmune diseases, hepatic and renal
CC   disease, inflammation, allergies, ischaemic shock, Alzheimer's and
CC   cognitive disorders, schizophrenia, restenosis, cardiovascular disorders,
CC   wound healing, stroke, arthritis, obesity, asthma, sepsis, acne,
CC   psoriasis, transplant rejection, metabolic disorders, infections and
CC   AIDS. The polypeptides are also useful for identifying their binding
CC   partners
XX
SQ   Sequence 43 AA;

  Query Match             100.0%;  Score 15;  DB 2;  Length 43;
  Best Local Similarity   60.0%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          4 RXLXF 8
              | | |
Db         28 RTLTF 32


RESULT 45
AAU86594
ID   AAU86594 standard; protein; 45 AA.
XX
AC   AAU86594;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Novel human connective tissue related polypeptide #160.
XX
KW   Human; connective tissue related disorder; cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200155343-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001322.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
```

```
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-565190/63.
DR   N-PSDB; ABK41772.
XX
PT   Nucleic acid encoding novel connective tissue associated polypeptides,
PT   used in diagnosing, preventing, treating or ameliorating a disorder such
PT   as cancer or rheumatoid arthritis.
XX
PS   Claim 11; SEQ ID NO 659; 673pp; English.
XX
CC   The present invention relates to the isolation of novel human connective
CC   tissue related polypeptides and the polynucleotide (cDNA and genomic)
CC   sequences encoding them. The sequences of the invention are useful in the
CC   diagnosis, treatment, prevention and/or prognosis of diseases associated
CC   with connective tissue(s), including cancer. The polynucleotide sequences
CC   of the invention are also useful in gene therapy. AAU86435-AAU86923
CC   represent the novel human connective tissue related polypeptides. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 45 AA;

  Query Match             100.0%;  Score 15;  DB 4;  Length 45;
  Best Local Similarity    60.0%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RXLXF 8
              | | |
Db         27 RALSF 31


RESULT 46
ADB59928
ID   ADB59928 standard; protein; 45 AA.
XX
AC   ADB59928;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Connective tissue antigen (CTA) #160.
XX
KW   cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
KW   antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;
KW   antiinflammatory; antiallergic; antiasthmatic; dermatological;
KW   nephrotopic; virucide; fungicide; antibacterial; antiparasitic;
KW   gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
KW   systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
KW   cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; cardiovascular disease;
KW   atherosclerosis; myocarditis; cardiopulmonary bypass complication;
KW   autoimmune disease; multiple sclerosis; allergic reaction; asthma;
KW   rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
KW   gastrointestinal disorder; inflammatory bowel disease;
KW   organ transplant rejection; immune system disorder; Bruton's disease;
KW   X-linked lymphoproliferative syndrome;
KW   B-cell lymphoproliferative disorder; HIV; AIDS; infection;
KW   chromosome identification; chromosome mapping; connective tissue antigen;
KW   CTA.
XX
OS   Homo sapiens.
XX
PN   US2003054375-A1.
XX
PD   20-MAR-2003.
XX
PF   07-MAR-2002; 2002US-00092154.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
```

PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764847.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-634869/60.
DR N-PSDB; ADB59439.
XX
PT New connective tissue-related polypeptides and polynucleotides, useful
PT for treating, preventing and/or prognosing e.g. disorders of connective
PT tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or
PT neoplasias.
XX
PS Claim 11; SEQ ID NO 659; 248pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I), which
CC comprises a sequence that is at least 95 % identical to a connective
CC tissue-related polynucleotide encoding connective tissue antigens (CTA).
CC The polypeptide or polynucleotide is useful for preventing, treating, or
CC ameliorating medical conditions in a mammal. The connective tissue
CC polypeptides, polynucleotides and antibodies are particularly useful for
CC treating, preventing and/or prognosing disorders of connective tissues
CC (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,
CC scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or
CC neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.
CC Alzheimer's disease, or Parkinson's disease), cardiovascular diseases
CC (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass
CC complications), autoimmune diseases (e.g. systemic lupus erythematosus,
CC rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.

Query Match 100.0%; Score 15; DB 7; Length 45;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 RXLXF 8
              | | |
Db         27 RALSF 31
```

RESULT 47
ADN42941
ID ADN42941 standard; protein; 45 AA.
XX
AC ADN42941;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human secreted protein SECP-2.
XX
KW cytostatic; anorectic; immunosuppressive; gene therapy; SECP-antagonist;
KW SECP-agonist; secreted protein; SECP; autoimmune disorder; obesity;
KW cancer; human; SECP-2.
XX
OS Homo sapiens.
XX
PN WO2004037987-A2.
XX
PD 06-MAY-2004.
XX
PF 22-OCT-2003; 2003WO-US033491.
XX
PR 22-OCT-2002; 2002US-0420720P.
PR 07-NOV-2002; 2002US-0425207P.
PR 15-NOV-2002; 2002US-0426679P.
PR 19-NOV-2002; 2002US-0427871P.
PR 06-JAN-2003; 2003US-0438551P.
PR 17-JAN-2003; 2003US-0441144P.
PR 27-JAN-2003; 2003US-0443135P.

XX
PA (INCY-) INCYTE CORP.
XX
PI Baughn MR, Becha SD, Bhatia UG, Blake JJ, Burrill JD, Chawla NK;
PI Chien D, Elliott VS, Emerling BM, Favero KD, Hafalia AJA;
PI Harmssen BW, Ho A, Ison CH, Khare R, Lee S, Lee SY, Lu DAM;
PI Marquis JP, Murage J, Nguyen DB, Ramkumar J, Richardson TW;
PI Swarnakar A, Tang TY, Tran UK, Wang JT, Yue H, Zheng W;
XX
DR WPI; 2004-365505/34.
DR N-PSDB; ADN43016.
XX
PT New human secreted protein (SECP) polypeptide, useful for preparing a
PT composition for treating a disease associated with decreased expression
PT or overexpression of functional SECP e.g., autoimmune disorders, obesity
PT or cancer.
XX
PS Claim 1; SEQ ID NO 2; 289pp; English.
XX
CC The invention describes an isolated human secreted protein (SECP)
CC polypeptide. Also described are: an isolated polynucleotide encoding the
CC polypeptide; a recombinant polynucleotide comprising a promoter sequence
CC operably linked to the polynucleotide; a cell transformed with the
CC recombinant polynucleotide; a transgenic organism comprising the
CC recombinant polynucleotide; a method of producing the polypeptide; an
CC isolated antibody that specifically binds to the polypeptide; a method of
CC detecting a target polynucleotide in a sample; a method for treating a
CC disease or condition associated with decreased expression or
CC overexpression of functional SECP; a method of screening a compound for
CC effectiveness as an agonist or antagonist of the polypeptide or in
CC altering expression of the target polynucleotide; a method of screening a
CC compound that specifically binds to, or that modulates the activity of,
CC the polypeptide; a method of assessing toxicity of a test compound; a
CC diagnostic test for a condition or disease associated with the expression
CC of SECP in a biological sample; a method of diagnosing a condition or
CC disease associated with the expression of SECP in a subject; a
CC composition comprising the antibody and a carrier, or the polypeptide, or
CC agonist or antagonist compound and an excipient a method of preparing a
CC polyclonal or monoclonal antibody; a method of detecting the polypeptide
CC in a sample; a method of purifying the polypeptide; a method of
CC generating an expression profile of a sample that contains
CC polynucleotides; and an array comprising different nucleotide molecules
CC affixed in distinct physical locations on a solid substrate, where at
CC least one of the nucleotide molecules comprises a first oligonucleotide
CC or polynucleotide sequence specifically hybridizable with at least 30
CC contiguous nucleotides of the target polynucleotide. The polypeptide is
CC useful for preparing a composition for diagnosing or treating a disease
CC or condition associated with decreased expression or overexpression of
CC functional SECP e.g. autoimmune disorders, obesity or cancer. This is the
CC amino acid sequence of a human secreted protein.
XX
SQ Sequence 45 AA;

Query Match 100.0%; Score 15; DB 8; Length 45;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 40 RALSF 44

RESULT 48
ABG26048
ID ABG26048 standard; protein; 49 AA.
XX
AC ABG26048;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #26039.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS90235.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 56407; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 49 AA;

Query Match 100.0%; Score 15; DB 4; Length 49;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 17 RALAF 21

RESULT 49
ADS98736
ID ADS98736 standard; protein; 49 AA.
XX
AC ADS98736;
XX
DT 30-DEC-2004 (first entry)
XX
DE Protein factor discovery related human contig polypeptide, SEQ ID 1000.
XX
KW antiinflammatory; cytostatic; antimicrobial; gene therapy; inflammation;
KW leukaemia; nervous system disorder; infection.
XX

OS Homo sapiens.
XX
PN WO2004087874-A2.
XX
PD 14-OCT-2004.
XX
PF 24-MAR-2004; 2004WO-US009202.
XX
PR 28-MAR-2003; 2003US-0458824P.
XX
PA (NUVE-) NUVELO INC.
PA (DRMA/) DRMANAC R T.
XX
PI Tang YT, Zhou P, Wang J, Wang ZW, Hu T;
XX
DR WPI; 2004-737686/72.
DR N-PSDB; ADS98396.
XX
PT New polynucleotides encoding a polypeptide with biological activity,
PT useful for treating inflammation, leukemias, nervous system disorders, or
PT infections.
XX
PS Example 3; SEQ ID NO 1000; 253pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising any
CC of the 235 nucleotide sequences described in the specification. The
CC invention further comprises: an isolated polynucleotide encoding a
CC polypeptide with biological activity, where the polynucleotide hybridizes
CC to one of the 235 novel polynucleotides under stringent hybridization
CC conditions, or having greater than about 99% sequence identity with the
CC novel polynucleotide; a vector comprising a novel polynucleotide; an
CC expression vector comprising the novel polynucleotide; a host cell
CC genetically engineered to comprise the novel polynucleotide, which can be
CC operatively associated with a regulatory sequence that modulates
CC expression of the polynucleotide in the host cell; an isolated
CC polypeptide encoded by the novel polynucleotide, or a polynucleotide
CC hybridizing under stringent conditions to the novel polynucleotide; a
CC composition comprising the polypeptide and a carrier; an antibody
CC directed against the polypeptide; a method for detecting the novel
CC polynucleotide in a sample; a method for detecting the polypeptide in a
CC sample; a method for identifying a compound that binds to the polypeptide
CC ; a method for producing the polypeptide; an isolated polypeptide
CC comprising any of the 235 amino acid sequences described in the
CC specification; and a collection of polynucleotides comprising of at least
CC one of the polynucleotides cited above. The polypeptides and
CC polynucleotides of the invention have antiinflammatory, cytostatic, and
CC antimicrobial activities. The novel polynucleotide may be used to treat
CC disorders by gene therapy. The polypeptides and polynucleotides are
CC useful for treating inflammation, leukaemias, nervous system disorders,
CC or infections. This sequence represents the polypeptide encoded by
CC contiguous DNA derived from one of the 235 novel isolated polynucleotides
CC of the invention.
XX
SQ Sequence 49 AA;

Query Match 100.0%; Score 15; DB 8; Length 49;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RXLXF 8
| | |
Db 17 RALAF 21

RESULT 50
AAM86216
ID AAM86216 standard; protein; 50 AA.
XX
AC AAM86216;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:13809.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
DR N-PSDB; AAK58997.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Claim 11; SEQ ID NO 13809; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 50 AA;

Query Match 100.0%; Score 15; DB 4; Length 50;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 RXLXF 8
              | | |
Db          6 RSLAF 10
```

Search completed: May 4, 2006, 06:05:50
Job time : 196 secs